

Grundlagen, Konzepte und neue Methoden der Zeitreihenanalyse

Stier, Winfried

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:

Stier, W. (1991). Grundlagen, Konzepte und neue Methoden der Zeitreihenanalyse. In H. Best, & H. Thome (Hrsg.), *Neue Methoden der Analyse historischer Daten* (S. 201-238). Sankt Katharinen: Scripta Mercaturae Verl. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-338098>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Grundlagen, Konzepte und neue Methoden der Zeitreihenanalyse

von Winfried Stier

Einleitung

Im folgenden sollen grundlegende Konzepte und Werkzeuge der Zeitreihenanalyse in möglichst einfacher Form dargestellt werden. Deshalb wird bewußt auf mathematische Ableitungen und Feinheiten verzichtet. Formeln werden nur insoweit eingeführt, als sie zum Verständnis unverzichtbar sind. Die Darstellung ist für Leser gedacht, die keine oder nur rudimentäre Vorkenntnisse in Zeitreihenanalyse besitzen. Selbstverständlich kann hier kein Gesamtüberblick über dieses mittlerweile riesige Teilgebiet der Statistik gegeben werden. Was jeweils über die Grundlagen hinaus zur Darstellung kommt, ist mehr oder weniger das Resultat subjektiver Auswahl des Autors und/oder der Herausgeber.

I. Traditionelle Zeitreihenmodelle

1. Grundlagen

Die Zeitreihenanalyse spielt seit jeher in der empirischen Wirtschaftsforschung eine bedeutende Rolle. Von daher stammt auch das sog. "Komponentenmodell", das für traditionelle Zeitreihenanalysen häufig die Grundlage bildet. Dabei wird unter einer Zeitreihe eine Folge von zeitlich geordneten, äquidistanten Beobachtungswerten eines mindestens auf Intervallskalenniveau gemessenen Merkmals verstanden. Grundlegend für das Komponentenmodell ist die Vorstellung, daß sich die beobachteten Reihenwerte aus verschiedenen "Bewegungskomponenten" zusammensetzen. Üblicherweise werden vier derartige Komponenten unterschieden:

a) Eine "Trendkomponente" T_t , deren Verlauf als durch langfristig wirkende Ursachen bedingt angesehen wird. Oft wird unterstellt, daß sie monoton wächst (z.B. auf Grund des technischen Fortschritts) oder monoton fällt (als Folge z.B. eines Bevölkerungsrückgangs).

b) Eine "Zyklische Komponente" Z_t , deren Verlauf den Konjunkturzyklus reflektiert und für die deshalb eine "wellenförmige" Bewegung postuliert wird.

c) Eine "Saisonkomponente" S_t , deren Verlauf auf jahreszeitliche und institutionelle Ursachen zurückgeführt wird. Für sie wird ebenfalls ein "wellenförmiger" Verlauf angenommen.

d) Und schließlich eine "irreguläre Komponente" U_t , deren Verlauf nicht auf die bei den anderen Komponenten aufgeführten Ursachenkomplexen zurückgeführt werden kann. Es wird angenommen, daß die U_t -Werte relativ (d.h. im Vergleich zu den T_t -, Z_t -, und den S_t -Werten) klein sind und "regellos", um den Wert Null schwanken. Die U_t sind Residualgrößen, die den Charakter von "Zufallsschwankungen" aufweisen sollen.

Häufig werden die Komponenten T_t und Z_t nicht getrennt betrachtet, sondern zur sog. "glatten Komponente", G_t , zusammengefaßt. Dies hängt nicht zuletzt damit zusammen, daß eine Trennung dieser beiden Komponenten mit traditionellen zeitreihenanalytischen Werkzeugen als äußerst problematisch anzusehen ist.

Die eben skizzierte inhaltliche Deutung der Komponenten reicht natürlich für eine Analyse nicht aus. Dazu ist mindestens eine Vorstellung darüber erforderlich, wie die einzelnen Komponenten im Zeitablauf zusammenwirken. Das einfachste Modell postuliert eine additive Überlagerung dieser Komponenten. Für die Zeitreihe X_t gilt mit diesem Postulat somit:

$$X_t = T_t + Z_t + S_t + U_t, \quad t = 1, 2, \dots, T$$

Der Index t bezieht sich dabei auf vorgegebene Zeitperioden, für die das Beobachtungsmaterial vorliegt. Diese sind meistens Monate, Quartale, Halbjahre. Aber auch Tageswerte (z.B. bei Börsenkursen) oder Wochenwerte finden sich in der Praxis. Für Jahreswerte hat die Komponente S_t natürlich keinen Sinn, da saisonale Schwankungen per Definition nur bei unterjährigen Daten existieren. Hier sind nur drei Komponenten sinnvoll.

Statt einer additiven Verknüpfung wird häufig von einer multiplikativen Komponentenverknüpfung ausgegangen, so z.B. bei einem Saisonbereinungsverfahren, das als "Census-X-11", bekannt ist. (Auf derartige Verfahren wird gleich noch näher einzugehen sein). Formal kann das multiplikative Modell durch Logarithmierung auf ein additives zurückgeführt werden. Neben den "reinen Typen" sind auch Mischformen denkbar, so z.B. wenn angenommen wird, daß sich die saisonale Komponente proportional zur glatten Komponente entwickelt.

2. Probleme der Zeitreihenanalyse

An obigem traditionellen Komponentenmodell lassen sich auf einfache Weise eine Reihe von Fragestellungen und Probleme diskutieren, die prinzipiell auch bei komplizierteren Zeitreihenansätzen in gleicher oder ganz ähnlicher Weise auftreten. Ein wichtiges praktisches Problem ist die sog. "Saisonbereinigung". Dabei soll die Komponente S_t identifiziert und eliminiert werden. Es ist unmittelbar einleuchtend, daß saisonbereinigte Reihen in der Praxis eine wichtige Rolle spielen. Ist z.B. eine Zunahme der Zahl der Arbeitslosen ein Indiz für einen sich tendenziell verschlechternden Arbeitsmarkt oder ist diese nur jahreszeitlich bedingt, also kurzfristiger Natur? Um diese und ähnliche Fragen beantworten zu können, muß eine Reihe um die Saisonkomponente "bereinigt" werden. Dies geschieht in der Hoffnung, daß mit Hilfe der bereinigten Reihe die Frage beantwortet werden kann, wie die Entwicklung verlaufen wäre, wenn eben keine jahreszeitlichen Einflüsse wirksam gewesen wären.

Eine andere Frage wäre z.B. die nach der "langfristigen" Entwicklung einer Reihe, d.h., wie würde sich eine Reihe ohne alle kurzfristigen Einflüsse entwickeln. Zur Beantwortung dieser Frage muß neben der saisonalen auch die irreguläre Komponente eliminiert werden. Diese Frage ist natürlich auch sinnvoll für Jahresreihen oder für Reihen mit mehrjährigem Rhythmus (z.B. Fünfjahresreihen), wie sie etwa in der historischen Forschung häufig vorkommen. Hier interessiert man sich z.B. für einen "säkularen" Trend. Oder dafür, ob eine Reihe eine zyklische Komponente enthält und falls ja, welche Zyklenlänge vorliegt. Offensichtlich ist das letztere Problem wesentlich komplizierter als die vorher erwähnten, denn die Isolation einer zyklischen Komponente setzt eine Elimination aller übrigen Komponenten voraus. Wesentlich ist dabei, daß deren Elimination in einer Weise erfolgt, daß die zyklische Komponente nicht tangiert wird. Auf diesen Problemkreis wird noch zurückzukommen sein.

Wie diese Ausführungen zeigen, besteht ein wichtiges Ziel der Zeitreihenanalyse darin, gewisse Komponenten einer Reihe zu isolieren oder zu eliminieren. Die Frage stellt sich deshalb, welche Werkzeuge die traditionelle Zeitreihenanalyse zur Verfügung stellt, um die angesprochenen Probleme zu lösen.

3. Werkzeuge der traditionellen Zeitreihenanalyse

In diesem Abschnitt sollen einige Werkzeuge der traditionellen Zeitreihenanalyse besprochen werden. Diese spielen durchaus auch heute noch eine bedeutende Rolle in der Praxis, allerdings häufig in modifizierter und/oder verfeinerter Form. Ihr Einsatz läßt sich am einfachsten im Kontext eines konkreten zeitreihenanalytischen Problems zeigen. Insbesondere können dabei sowohl theoretische als auch praktische Probleme, die mit ihrem Einsatz verknüpft sind, anschaulich diskutiert werden.

Ausgegangen werden soll vom oben schon ansatzweise diskutierten Problem der Saisonbereinigung, wobei das traditionelle, additive Komponentenmodell zugrundegelegt wird. Ohne Beschränkung der Allgemeinheit wird von monatlichen Reihen ausgegangen. Dieses Komponentenmodell allein reicht jedoch nicht aus, um die gestellte Aufgabe zu lösen. Vielmehr ist es erforderlich, für die einzelnen Komponenten bestimmte Verlaufshypothesen zu formulieren:

Es sei angenommen, daß

$$a) \quad S_t = S_{t+12}, \quad t=1,2,\dots,T$$

ist, d.h. eine streng periodische Saisonfigur wird unterstellt.

b) die glatte Komponente innerhalb eines Zeitraumes von 13 Monaten durch eine lineare Funktion der Zeit hinreichend genau approximiert werden kann.

Unter diesen beiden Voraussetzungen und der weiteren, daß die Saisonfigur auf Null normiert wird, d.h., daß gilt $S_1 + S_2 + \dots + S_{12} = 0$, kann die glatte Komponente mit Hilfe eines gleitenden symmetrischen 12-Monats-Durchschnitts "geschätzt" werden:

$$\hat{G} = \frac{1}{12} \left[\frac{1}{2} X_{t-1} + \frac{1}{2} X_{t+1} + \sum_{i=t-5}^{t+5} X_i \right]$$

Die Normierung der Saisonfigur auf die Summe Null bedeutet keine Einschränkung der Allgemeinheit des Modellansatzes, da ja im Fall einer von Null verschiedenen Summe eine Konstante von den S-Werten subtrahiert und der glatten Komponente zugeschlagen werden könnte.

Problematisch sind allerdings die Postulate a) und b). Abgesehen davon, daß die postulierte strenge Periodizität von S_t völlig unrealistisch ist (wie man durch Betrachten konkreter Reihen leicht erkennt), bleibt anzumerken, daß das benutzte Werkzeug zur Isolation der glatten Komponente eine ganz bestimmte Verlaufsform für G_t impliziert. Es kann nämlich gezeigt werden, daß dieser gleitende Durchschnitt die glatte Komponente nur dann fehlerfrei erfaßt, wenn diese tatsächlich linear im angegebenen "Stützbereich" von 13 Monaten verläuft.

Dieses Beispiel zeigt sehr anschaulich, daß das verwendete Instrumentarium (gleitende Durchschnitte) eigentlich eine dazu "passende" Definition der zu erfassenden Komponente voraussetzt. Es wird zwar keine "globale" Linearität (d.h. für alle Zeitpunkte) vorausgesetzt, wohl aber eine (weniger rigide) "lokale" Linearität (d.h. für jeweils 13 Monate). Verläuft nun G_t nicht linear innerhalb der 13 Monate-Stützbereiche, dann kann die glatte Komponente offensichtlich nicht "sauber" isoliert werden. Die Fehler sind umso gravierender, je weniger die Verlaufshypothese zutrifft. Deswegen wurde in b) schon vorsichtig von einer "hinreichenden" Genauigkeit gesprochen, ohne daß diese allerdings präzisiert wurde.

Methodologisch unbefriedigend ist insbesondere der Sachverhalt, daß bei diesem Verfahren keine operationale Definition der glatten Komponente a priori gegeben wird, diese sich vielmehr a posteriori aus dem verwendeten Analysewerkzeug ergibt. Eigentlich müßte es ja gerade umgekehrt sein: das Analysewerkzeug müßte so konstruiert werden, daß es in "optimaler" Weise der auf substanzwissenschaftlichen Überlegungen beruhenden Komponentendefinition gerecht wird. Diese Problematik wird unten nochmals zu diskutieren sein.

Im Vorgriff auf spätere Darlegungen sei hier schon erwähnt, daß gleitende Durchschnitte als "Filter" interpretiert werden können. Mit Hilfe filtertheoretischer Methoden kann gezeigt werden, daß symmetrische gleitende Durchschnitte die Eigenschaft der sog. "Nullphasigkeit" haben, d.h. die ausgefilterten Werte werden den richtigen Zeitpunkten zugeordnet, so daß sich keine Verzögerungen zwischen der Originalreihe und der glatten Komponente ergeben. Dabei ist jedoch auch sogleich auf einen praktisch gravierenden Nachteil gleitender Durchschnitte aufmerksam zu machen: sowohl am Reihenanfang als auch am Reihende gehen Werte verloren (bei obigem gleitenden Durchschnitt je 6). Dies impliziert, daß die glatte Komponente insbesondere nicht bis zum "aktuellen Rand" einer Reihe bestimmt werden kann. Da sich das Interesse aber in der Praxis gerade häufig auf diesen

Randbereich konzentriert (Verwendung bereinigter Reihen zu Diagnosezwecken), ist dieser Sachverhalt besonders schwerwiegend. Man könnte deshalb auf die Idee kommen, asymmetrische gleitende Durchschnitte zu verwenden, die außer dem jeweiligen aktuellen Reihenwert nur vergangene Reihenwerte verarbeiten. Formal kommt man damit bis zum aktuellen Rand, jedoch weisen solche gleitende Durchschnitte eine (negative) Phase auf, d.h. der letzte gefilterte Wert ist nicht dem letzten Zeitpunkt zuzuordnen, sondern einem früheren. Je nach Konstruktion eines asymmetrischen gleitenden Durchschnitts ist das Ausmaß dieser Phasenverschiebung mehr oder weniger praktisch bedeutsam. Die sich hier abzeichnende "Randproblematik" tritt nicht nur bei traditionellen Werkzeugen der Zeitreihenanalyse auf.

Die saisonbereinigte Reihe ergibt sich schließlich mit Hilfe einiger einfacher Operationen. Da \hat{G}_t eine "Schätzung" für G_t ist, kann geschrieben werden:

$$X_t - \hat{G}_t \approx S_t + U_t.$$

Die $S_t + U_t$ - Werte werden nun über gleichnamige Monate gemittelt. Dazu ist eine Doppelindizierung zweckmäßig: X_{ij} sei der Reihenwert des j -ten Monats im i -ten Jahr und $S_{ij} = S_j$ auf Grund der postulierten Periodizität. Dann ist:

$$\tilde{S}_j = \frac{1}{n_j} \sum_i (X_{ij} - \hat{G}_{ij}) \approx S_{ij} + \frac{1}{n_j} \sum_i U_{ij}, \quad j = 1, 2, \dots, 12$$

n_j bedeutet die Anzahl der Jahre für die S_j berechenbar ist. Die Normierung der 12 Werte S_1, S_2, \dots, S_{12} erfolgt durch Subtraktion ihres arithmetischen Mittels:

$$\hat{S}_j = \tilde{S}_j - \frac{1}{12} \sum_{j=1}^{12} \tilde{S}_j, \quad j = 1, 2, \dots, 12$$

Die saisonbereinigte Reihe ergibt sich schließlich durch Bildung der Differenzen $X_{ij} - \hat{S}_j$.

Schließlich sei noch angemerkt, daß dieses einfache Verfahren in den dreißiger Jahren vom Federal Reserve Board (USA) und vom Londoner Wirtschaftsdienst für praktische Analysen verwendet wurde.

Später entwickelte und heute gebräuchliche Saisonbereinigungsverfahren stellen natürlich vergleichsweise wesentlich raffiniertere und kompliziertere

Konstruktionen dar, obwohl sie in der Regel immer noch auf dem ursprünglichen Komponentenmodell beruhen. Diese Konstruktionen unterscheiden sich vom obigen einfachen (und ähnlichen hier nicht besprochenen) Verfahren in verschiedener Hinsicht:

a) Zum einen werden a priori bestimmte Verlaufshypothesen für die einzelnen Komponenten eingeführt, die als flexibler bezeichnet werden können und für die dann jeweils geeignete Schätzverfahren verwendet oder entwickelt werden. Wenn hier von Schätzverfahren gesprochen wird, dann ist damit gemeint, daß es sich um Prozeduren im Sinne der statistischen Schätztheorie handelt. Dies setzt voraus, daß für das traditionelle Zeitreihenmodell eine "geeignete" stochastische Spezifikation formuliert wird, was in der Regel dadurch erfolgt, daß U_t nunmehr als eine Folge von Zufallsvariablen (d.h. als stochastischer Prozeß, dazu siehe unten) interpretiert wird. Sowohl die glatte als auch die saisonale Komponente wird nach wie vor als deterministische Funktion der Zeit aufgefaßt (über Ausnahmen siehe unten).

b) Zum anderen wird neben gleitenden Durchschnitten vor allem die Methode der Kleinsten-Quadrate verwendet. Diese Methode ist eine in der Statistik generell häufig angewandte Schätzprozedur. Sie beruht auf einem einfachen Grundgedanken, der an Hand des folgenden bewußt simplen Zeitreihenmodelles illustriert werden soll. Angenommen, es gelte das Zwei-Komponentenmodell:

$$X_t = G_t + U_t, t=1,2, \dots, T$$

und für die glatte Komponente werde eine lineare Verlaufsform angenommen, d.h. $G_t = a + bt, t=1,2,\dots,T$, wobei die unbekannten Parameter a und b aus der vorliegenden Zeitreihe zu bestimmen sind. Die irreguläre Komponente soll zunächst obigen - etwas vage formulierten - Annahmen genügen. Nach dem Kleinst-Quadrate-Prinzip werden nun die Parameter so bestimmt, daß die Summe der "Abweichungsquadrate"

$$\sum_{t=1}^T (X_t - G_t)^2 = \sum_{t=1}^T (X_t - \hat{a} - \hat{b}t)^2, \quad t = 1, 2, \dots, T$$

minimiert wird bezüglich \hat{a} und \hat{b} . \hat{a} und \hat{b} sind die Kleinst-Quadrate-"Schätzungen" von a und b . Sie lassen sich technisch durch partielle Differentiation bestimmen.

Hier ist nun eine Zwischenbemerkung zu machen. Bei dieser Kleinst-Quadrate-Prozedur ist von einem sog. "deskriptiven" Zeitreihenmodell ausge-

gangen worden, d.h. U_t wurde einfach als eine "Rest"-Komponente mit gewissen qualitativ umschriebenen Eigenschaften angesehen, die sie als "zufällig" qualifizieren soll. Präzisiert man jedoch "zufällig" im Sinne der Wahrscheinlichkeitstheorie, dann erhalten die U_t den Status von Zufallsvariablen (genauer: einer Folge von Zufallsvariablen, vgl. unten) mit genau definierten Eigenschaften, was man mit "stochastischer Spezifikation" bezeichnet. Im einfachsten Fall besteht diese in den drei Annahmen:

- aa) die Mittelwerte (genauer: Erwartungswerte) der U_t sind gleich Null für alle $t=1,2,\dots,T$
- bb) die Varianzen der U_t sind konstant für alle $t=1,2,\dots,T$
- cc) die U_t sind unkorreliert.

Mit dieser stochastischen Spezifikation sind \hat{a} und \hat{b} Schätzfunktionen im Sinne der statistischen Schätztheorie, d.h. Zufallsvariablen. Die konkreten numerischen Werte für \hat{a} und \hat{b} werden als Realisationen dieser Zufallsvariablen betrachtet. Erst auf dieser Ebene können dann Aussagen über etwaige Optimalitätseigenschaften von Schätzverfahren gemacht werden. So kann z.B. gezeigt werden, daß \hat{a} und \hat{b} bei der angenommenen stochastischen Spezifikation BLUE (Best Linear Unbiased Estimates) sind, d.h. die unbekannten Parameter a und b werden "unverzerrt" mit kleinster Varianz geschätzt, wobei die Schätzfunktionen linear von den Beobachtungswerten X_t abhängen.

Schließlich sei noch angemerkt, daß zwischen gleitenden Durchschnitten und Kleinst-Quadrate-Schätzungen manchmal bestimmte Zusammenhänge bestehen. Bei gewissen Schätzproblemen stellen sich beide Ansätze als formal äquivalent heraus, d.h. eine Kleinst-Quadrate-Lösung läßt sich als gleitender Durchschnitt darstellen.

Eine Skizzierung einiger praktischer Verfahren soll das Gesagte illustrieren. Beim "Berliner Verfahren" z.B., dem offiziellen Saisonbereinungsverfahren des Statistischen Bundesamtes, wird davon ausgegangen, daß sich die glatte Komponente als Polynom dritter Ordnung der Zeit darstellen läßt. Eine Trennung von Trend und Zyklus erfolgt also nicht. Daß kein geringerer Polynomgrad gewählt wird, hängt damit zusammen, daß sonst keine konjunkturellen Wendepunkte modelliert werden könnten. Die Wahl des Polynomgrades ist also mindestens teilweise substanzwissenschaftlich motiviert. Weiter wird angenommen, daß die saisonale Komponente als trigonometrisches Polynom der Zeit formulierbar ist mit den "saisonalen Frequenzen" $\lambda_k = 2\pi k/12$, $k=1,\dots,6$. Für die irreguläre Komponente wird ein schwach

stationärer Prozeß (zu diesen Prozessen siehe unten) angenommen mit Erwartungswert Null. Für eine Zeitreihe kann dann geschrieben werden:

$$X_t = \sum_{j=1}^3 a_j t^j + \sum_{k=1}^6 (b_k \cos \lambda_k t + c_k \sin \lambda_k t) + u_t$$

Charakteristisch für das Verfahren ist nun, daß die Gültigkeit dieses Modelles nicht für alle Zeitpunkte $t = 1, 2, \dots, T$ angenommen wird, sondern nur für die Länge gewisser Stützbereiche welche das Intervall $[1, T]$ gleitend überdecken (sog. "gleitende Regression"). Die unbekannten Parameter des Modelles werden nun mit Hilfe der Methode der Kleinsten Quadrate bestimmt. Es kann gezeigt werden, daß die Kleinst-Quadrate-Lösung als gleitender Durchschnitt dargestellt werden kann (wobei die "Gewichte" ein komplizierteres Verlaufsmuster aufweisen als beim einfachen 12-er Durchschnitt), d.h. wiederum als spezieller Filter. Die Methode der kleinsten Quadrate kann in diesem Zusammenhang deshalb auch als eine - wenn auch sehr "indirekte" - Methode des Filter-Designs interpretiert werden (vgl. dazu die Ausführungen zur Filtertheorie unten).

Ein Bereinigungsverfahren, das im wesentlichen aus einer Abfolge von gleitenden Durchschnitten besteht, ist das Census X-11 -Verfahren, das z.B. von der Deutschen Bundesbank offiziell verwendet wird. Neben dem schon bekannten 12-er Durchschnitt werden 5-er, 9-er, 23-er usw. Durchschnitte in verschiedenen Phasen des iterativen Bereinigungsprozesses verwendet.

Neuere Verfahren, wie z.B. von Schlicht und Pauly-Schlicht nehmen für die glatte und die saisonale Komponente keine deterministische Funktion der Zeit mehr an, sondern wählen auch für diese Komponenten einen stochastischen Ansatz. Grundlegend ist aber auch hier das traditionelle Komponentenmodell.

Es kann hier nicht darum gehen, einen Überblick über Saisonbereinigungsverfahren zu geben. Der daran interessierte Leser sei dafür auf Stier (1985) verwiesen. Die angeführten Beispiele zeigen allerdings die Grundproblematik des traditionellen Komponentenmodelles sehr deutlich:

a) Für die einzelnen Komponenten muß eine ganz bestimmte Verlaufsform postuliert werden, egal ob der Ansatz deterministisch oder stochastisch ist.

b) Da es eine Vielzahl von denkbaren Verlaufsformen gibt, stellt sich die Frage nach der jeweils "geeigneten". Diese Frage ist aber von den Daten her

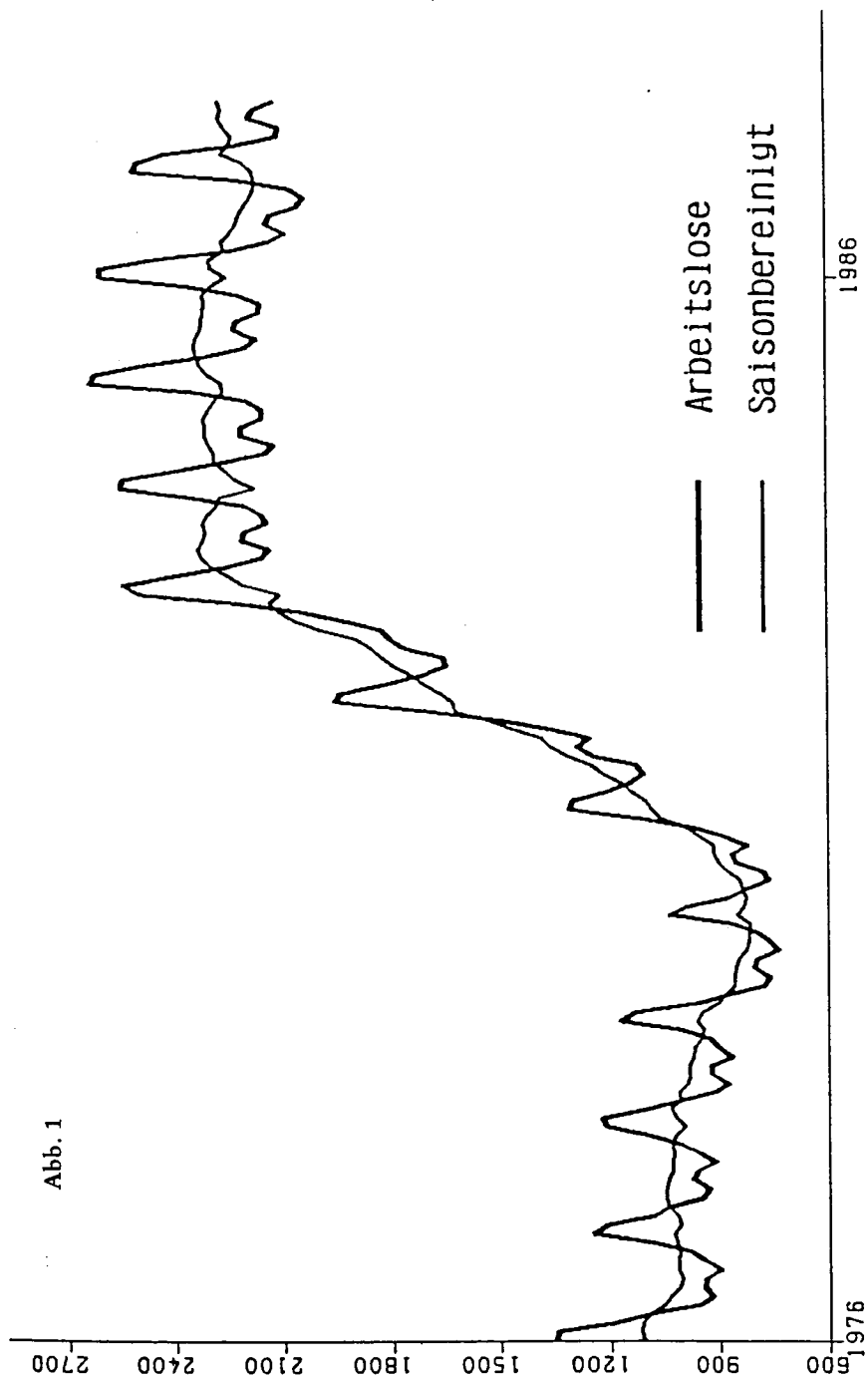
grundsätzlich nicht zu entscheiden, denn offensichtlich "funktionieren" die erwähnten (und auch die nicht erwähnten) Bereinigungsverfahren auf "beliebigen" Zeitreihen. Ein wertender Vergleich der einzelnen Verfahren dahingehend, welches "am besten" bereinigt, ist sehr problematisch, da die zu verwendenden Vergleichskriterien selbst nicht unumstritten sind.

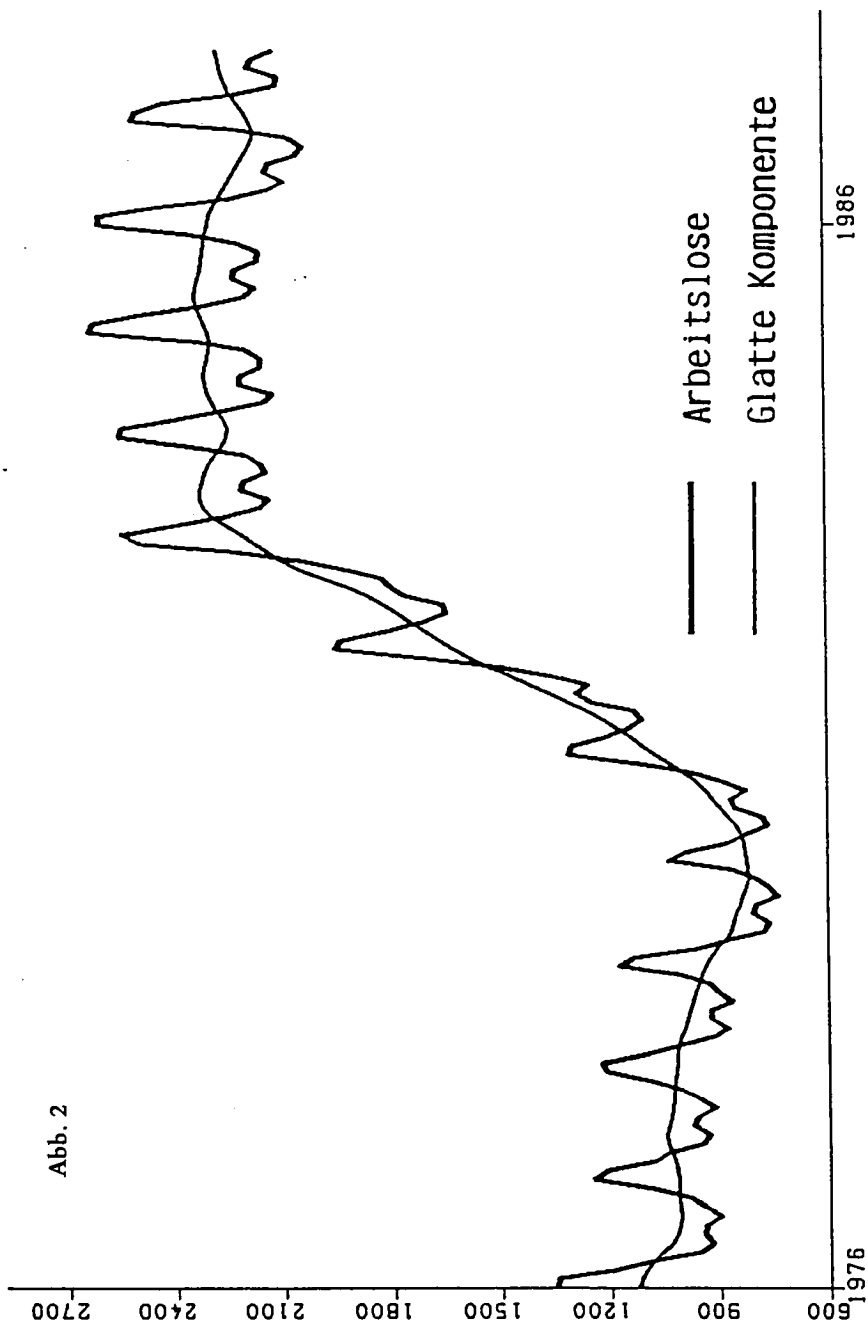
Schließlich seien hier die Resultate einer praktischen Saisonbereinigung graphisch wiedergegeben. Die monatliche Reihe "Arbeitslose in der BRD" (1976-1987) wurde mit Census X-11 (multiplikativ) bereinigt. Die beiden nachstehenden Abbildungen zeigen die saisonbereinigte Reihe (Abb. 1) und die glatte Komponente (Abb. 2), jeweils zusammen mit der Originalreihe.

4. Weitere Bemerkungen zum traditionellen Komponentenmodell

Die Ausführungen zum traditionellen Komponentenmodell in der Zeitreihenanalyse, die oben im Kontext der Saisonbereinigung gemacht wurden, gelten auch - unter jeweils zu beachtenden Modifikationen, Einschränkungen bzw. Erweiterungen - für andere Problemstellungen. Insbesondere treten auch hier die eben erwähnten Probleme a) und b) auf. Diese kann man zusammenfassend dahingehend charakterisieren, daß bei diesem Ansatz mit nicht-testbaren Hypothesen gearbeitet wird, was methodologisch ein bemerkenswertes Faktum ist, das jedoch häufig übersehen wird. Um die damit verbundenen Probleme nochmals zu verdeutlichen, sei hier abschließend kurz auf einige Aspekte der sog. "Langen-Wellen"-Forschung eingegangen, die meistens auf der Basis des traditionellen Komponentenmodelles betrieben wird.

Das zu lösende Grundproblem kann kurz so umrissen werden: Für viele historische Zeitreihen wird die Hypothese der Existenz langfristiger zyklischer Bewegungsabläufe als sinnvoll angesehen. Darunter werden quasi-periodische Abläufe verstanden mit einer Dauer von z.B. etwa 50 Jahren, um den längsten (nach Kondratieff benannten) Zyklus beispielhaft herauszugreifen. Daneben werden von einzelnen Forschern kürzere Zyklen von etwa 20 Jahren usw. postuliert. Einzelheiten interessieren hier nicht. Nun ist es bei historischen Reihen praktisch nie möglich, die Existenz derartiger Zyklen unmittelbar nachzuweisen, z.B. durch eine optische Inspektion einer Reihe, da andere, nicht-zyklische Komponenten, wie vor allem der Trend sowie "noise", dominieren. Anders ausgedrückt: denkt man sich die Varianz einer Reihe auf einzelne Reihenkomponten aufgeteilt, so entfällt nur ein geringer Teil dieser Varianz auf eine solche möglicherweise vorhandene zyklische Komponente.





Die Suche nach einer (oder mehreren) derartigen zyklischen Komponenten gleicht somit der Suche nach einer Nähnadel im Heuhaufen. Und sie kann auch nur dann mit Aussicht auf Erfolg unternommen werden, wenn es gelingt, die nicht-interessierenden Komponenten zu eliminieren. Dabei muß jedoch unbedingt sichergestellt werden, daß die verwendete Eliminationsprozedur weder vorhandene zyklische Komponenten tangiert (also z.B. ihre zyklischen Struktur verändert oder sie ganz oder teilweise eliminiert) noch Zyklen erzeugt. In der Regel besteht der erste Schritt darin, einen vorhandenen Trend zu eliminieren. Dabei stellt sich die Frage nach einer "geeigneten" Trendfunktion, wobei die oben diskutierten Probleme auftreten. Die Wahl einer speziellen Trendfunktion ist jedoch präjudizierend, denn der Verlauf der trendfreien Reihe hängt ganz entscheidend von dieser Wahl ab. Es leuchtet unmittelbar ein, daß die Suche nach zyklischen Komponenten in einer trendfreien Reihe ganz verschiedene Resultate erbringt, je nachdem, welcher Trend eliminiert wurde. Auf der Basis des klassischen Komponentenmodelles erscheint daher eine objektive Suche nach "langen Wellen" als ein aussichtsloses Unterfangen.

Wenn man von der Trendproblematik einmal absieht, verbleibt das weitere Problem der Elimination von kurzfristigen Schwankungen. Üblicherweise werden dazu gleitende Durchschnitte verwendet. Aber dabei stellt sich das Problem nach der Länge dieser Durchschnitte. Der Verlauf der geglätteten Reihe, die ja die zyklische Komponente in nuce darstellen soll, hängt nämlich davon ab. Außerdem ist bei Verwendung von gleitenden Durchschnitten, insbesondere wenn sie iterativ angewendet werden, darauf zu achten, daß durch sie keine Zyklen erzeugt werden. Es kann gezeigt werden, daß man durch Glätten von "noise" zyklische Reihen erzeugen kann (sog. Slutzy-Effekt). Weiter unten wird skizziert, wie man durch den Entwurf geeigneter Filter die erwähnten Probleme lösen bzw. umgehen kann.

II. Zeitreihenmodelle auf der Basis von stochastischen Prozessen

1. Grundlagen

Beim traditionellen Komponentenmodell wird unterstellt, daß sich eine Zeitreihe aus drei deterministischen Funktionen und einem Zufallsterm zusammensetzt. Der Zufall spielt in diesem Ansatz sozusagen die Rolle einer sachlich nicht weiter zu erklärenden Residualgröße. Bei neueren Modellkonzeptionen spielt der Zufall jedoch nicht mehr eine nur untergeordnete Nebenrolle. Vielmehr versucht man Zeitreihenmodelle zu bauen, die auf der

Theorie der stochastischen Prozesse beruhen. Unter einem stochastischen Prozeß versteht man - vereinfacht ausgedrückt - eine Folge von Zufallsvariablen, die im allgemeinen korreliert sind, also eine zeitliche Struktur aufweisen. Unter den stochastischen Prozessen sind die sog. "stationären" Prozesse mathematisch am weitesten entwickelt, insbesondere die sog. "schwach stationären" Prozesse. Diese sind einmal dadurch charakterisiert, daß sowohl ihre Mittelwert als auch ihre Varianzfunktion zeitinvariant sind. Zum anderen ist ihre Autokovarianz- bzw. Autokorrelationsfunktion nur vom "lag", d.h. der zeitlichen Distanz zweier Zufallsvariablen, abhängig und nicht von den jeweiligen absoluten Zeitpunkten.

Autokovarianz- und Autokorrelationsfunktion enthalten praktisch die gleichen Informationen über die zeitliche Struktur eines Prozesses. Während jedoch die Kovarianzfunktion vom verwendeten Maßstab abhängig und unbeschränkt ist, nimmt die Korrelationsfunktion nur Werte an, die absolut im Intervall $[0,1]$ liegen. Sie geht aus der Kovarianzfunktion durch eine geeignete Normierung hervor. Ein schwach stationärer Prozeß ist also durch die folgenden drei Postulate charakterisiert:

$$a) E(X_t) = m, t=1,2,\dots$$

$$b) \text{Var}(X_t) = \sigma^2, t=1,2,\dots$$

$$c) \text{Cov}(X_t, X_{t+s}) = E[(X_t - m)(X_{t+s} - m)] = C(s) \\ s=0,1,2,\dots, t=1,2,\dots$$

wobei s den zeitlichen Abstand je zweier Zufallsvariablen bezeichnet. Die Autokorrelationsfunktion $R(s)$ ist dabei durch

$$R(s) = C(s)/\text{Var}(X_t), s=0,1,2,\dots$$

$$\text{mit } |R(s)| \leq 1, \quad s = 0,1,2,\dots \text{ gegeben.}$$

Offensichtlich ist die oben charakterisierte Zufallskomponente U_t im klassischen Komponentenmodell nichts anderes als ein schwach stationärer Prozeß mit einer speziellen Korrelationsfunktion: die paarweisen Korrelationen der U_t , $t=1,2,\dots$ sind gleich Null. Einen derartigen Prozeß nennt man auch "weißes Rauschen" ("white noise"). Dieser einfachste Prozeßtyp erweist sich für viele Analysesituationen als grundlegend.

Für kompliziertere Prozesse sind natürlich auch kompliziertere Autokorrelationsfunktionen zu erwarten.

Die Entwicklung eines Zeitreihenmodelles mit Hilfe stochastischer Prozesse impliziert die Hypothese, daß eine gegebene Reihe als eine Realisation eines derartigen (wenn auch zunächst noch unbekannten) Prozesses interpretiert werden kann. Anders ausgedrückt: man versucht, von einer vorliegenden Reihe auf den "erzeugenden" Prozeß zu schließen, d.h. den "zugrundeliegenden" Prozeß zu "identifizieren". Dabei verwendet man aus guten Gründen möglichst einfache Prozeßtypen, deren Korrelationsverhalten theoretisch bekannt ist und für deren unbekannte Parameter Schätzprozeduren verfügbar sind.

2. AR-, MA-, ARMA- und ARIMA-Prozesse

Grundlegende "Bausteine" zur Modellbildung sind die autoregressiven und die gleitenden Durchschnittsprozesse. Die Prozeßgleichung für einen autoregressiven Prozeß der Ordnung p (kurz $AR(p)$ -Prozeß) lautet:

$$X_t - \Phi_1 X_{t-1} - \Phi_2 X_{t-2} - \dots - \Phi_p X_{t-p} = u_t, t=1,2,\dots$$

Dabei bezeichnet u_t weißes Rauschen. Damit $AR(p)$ ein stationärer Prozeß ist, müssen die Parameter $\Phi_1, \Phi_2, \dots, \Phi_p$, bestimmten Bedingungen genügen, auf die hier allgemein nicht eingegangen werden kann. Für $p=1$ z.B. muß Φ_1 betragsmäßig kleiner als Eins sein. Man sieht leicht, daß für $|\Phi_1| > 1$, der Prozeß "explodieren" würde. Die "mildeste" Form von Nicht-Stationarität liegt hier vor für $\Phi_1 = 1$. Die beiden nachstehenden Graphiken zeigen je eine Prozeßrealisierung von 200 Werten eines $AR(1)$ -Prozesses mit $\Phi_1 = 0.8$ und $\Phi_1 = 1$.

Interessant ist, daß sich das Niveau des Prozesses für $|\Phi_1| = 1$ dauernd ändert und tendenziell länger andauernde "trendartige" Auf- bzw. Abwärtsbewegungen zu beobachten sind. Dieser Prozeß wird auch als "random-walk" bezeichnet. Der Verlauf von Aktienkursen läßt sich z.B. häufig durch einen random-walk beschreiben.

Ein gleitender Durchschnittsprozeß der Ordnung q (auch "moving-average" - oder kurz $MA(q)$ -Prozeß genannt) ist wie folgt definiert:

$$X_t = u_t - \Theta_1 u_{t-1} - \Theta_2 u_{t-2} - \dots - \Theta_q u_{t-q}$$

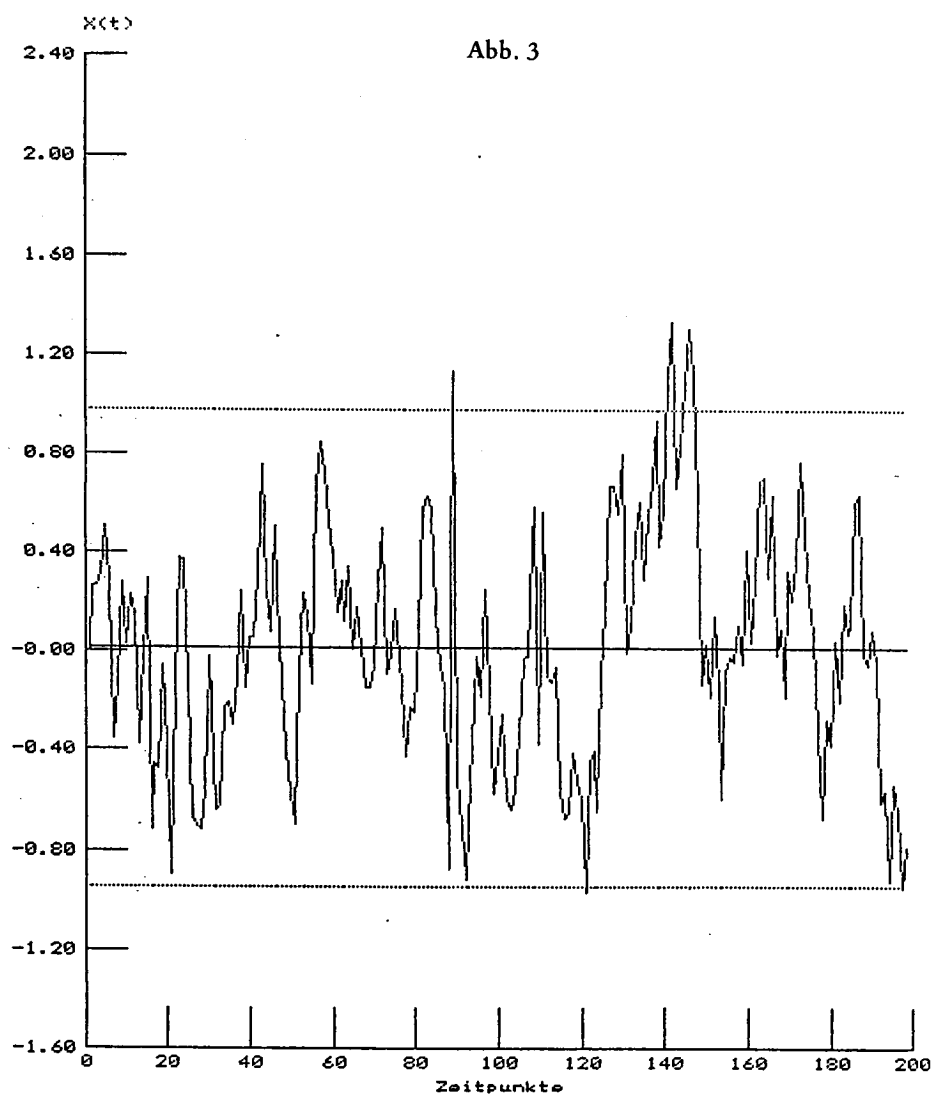
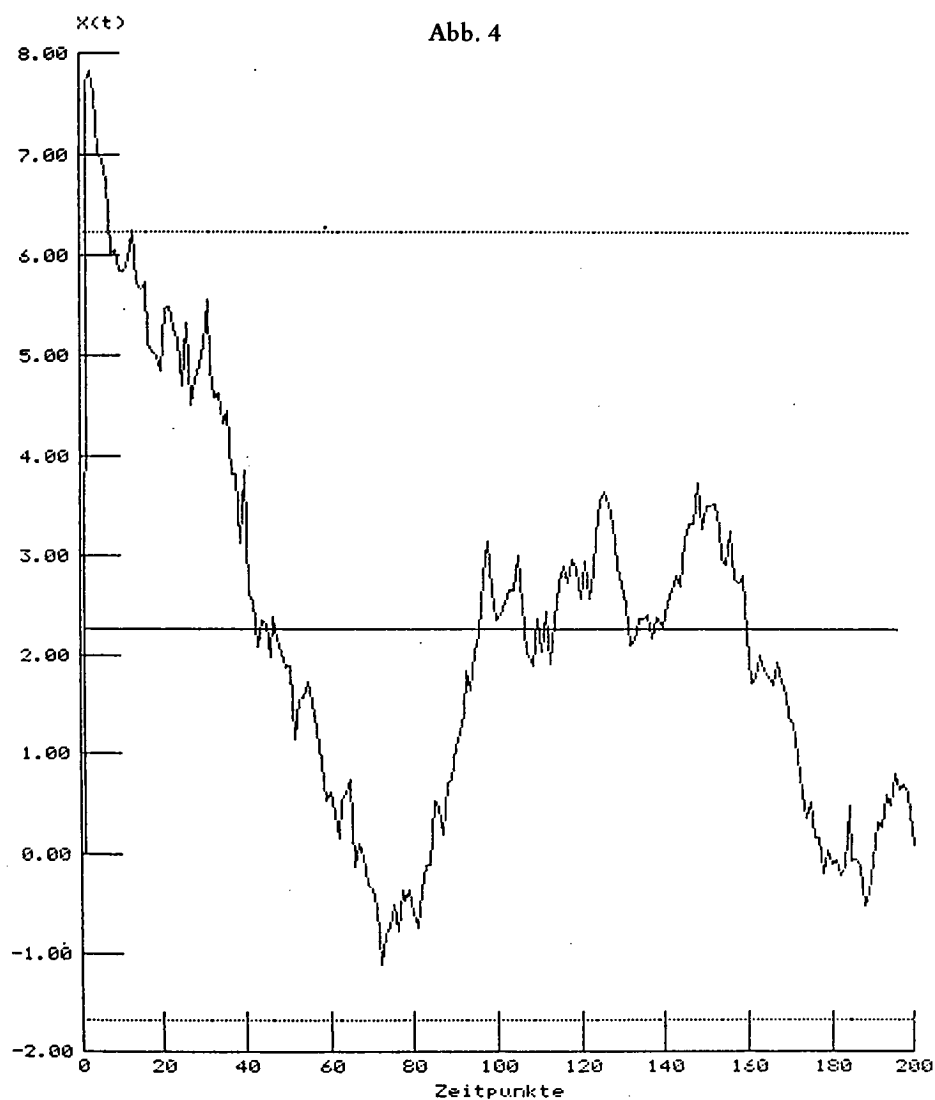
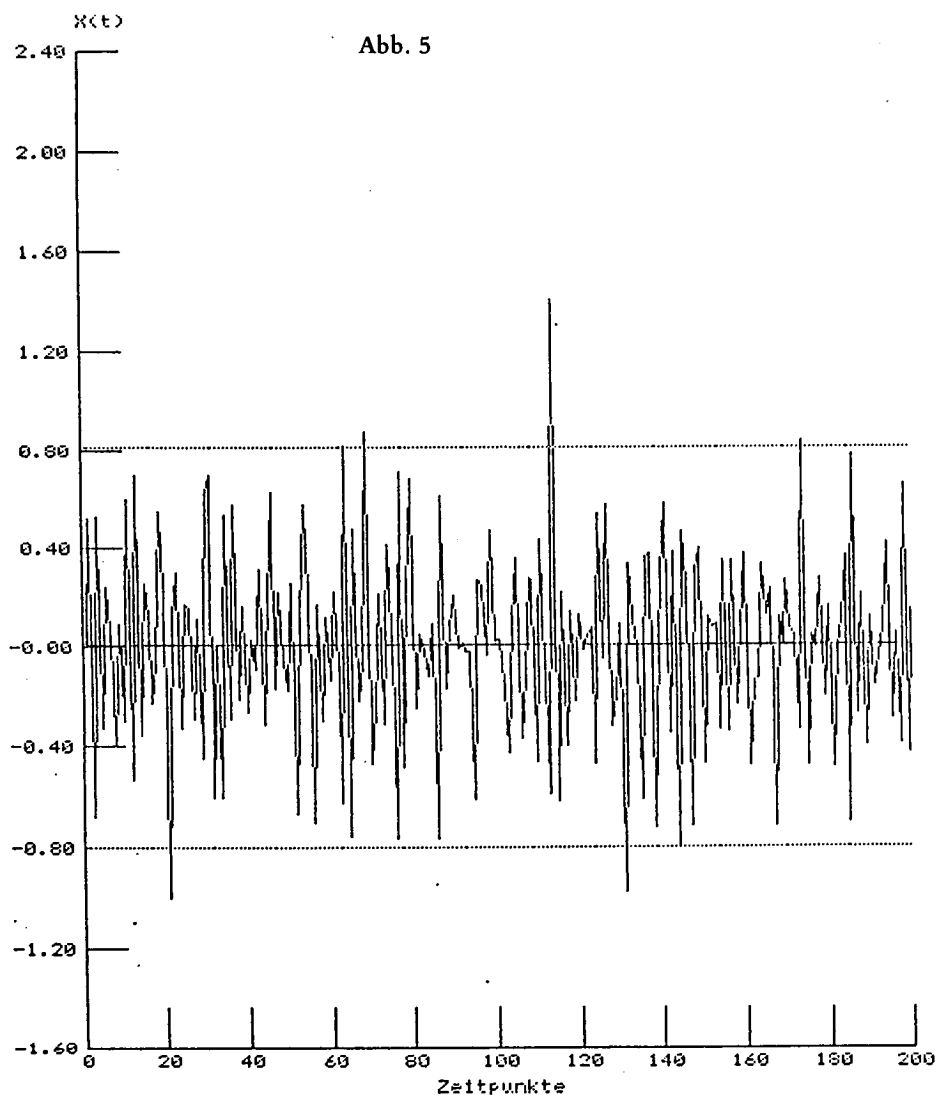


Abb. 4



wobei u_t wiederum weißes Rauschen bezeichnet. Nachstehende Abbildung zeigt eine Realisation eines MA(2)-Prozesses mit den Parametern $\Theta_1 = 0.9$ und $\Theta_2 = -0.1$.



Sowohl AR-als auch MA-Prozesse lassen sich mit Hilfe des "Verschiebungsoperators "

$$B^m(X_t) = X_{t-m}, \quad m=1,2,\dots$$

kürzer darstellen, wenn man damit die beiden " lag-Polynome"

$$\Phi(B) := 1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p$$

$$\Theta(B) := 1 - \Theta_1 B - \Theta_2 B^2 - \dots - \Theta_q B^q$$

eingführt. Die beiden Prozesse lassen sich dann kurz schreiben als

$$\Phi(B)X_t = u_t$$

$$\text{bzw. } X_t = \Theta(B)u_t$$

Bei der Modellierung von Zeitreihen zeigt sich nun häufig, daß eine Verwendung eines "reinen" Prozeßtyps, also entweder eines AR- oder eines MA-Prozesses, zu hohen Prozeßordnungen (d. h. zu einem großen p bzw. q) führen kann. Da in der Praxis die Prozeßparameter unbekannt sind, müssen sie aus dem vorliegenden Datenmaterial (d. h. einer endlichen Reihe) geschätzt werden. Je mehr Parameter geschätzt werden müssen, umso ungenauer fällt ihre Schätzung aus, d. h. umso größer ist ihre Varianz. Es empfiehlt sich daher, Modelle mit möglichst wenigen Parametern zu verwenden (sog. Prinzip der "parsimony"). Dies kann häufig durch eine Kombination beider Prozeßtypen erreicht werden. Solche werden als ARMA-Prozesse bezeichnet. Die Prozeßgleichung eines ARMA (p, q) -Prozesses ist gegeben durch:

$$X_t - \Phi_1 X_{t-1} - \Phi_2 X_{t-2} - \dots - \Phi_p X_{t-p} = u_t - \Theta_1 u_{t-1} - \Theta_2 u_{t-2} - \dots - \Theta_q u_{t-q},$$

oder kurz durch

$$\Phi(B)X_t = \Theta(B)u_t.$$

Zur Modellierung praktischer Zeitreihen wie sie im Bereich der Ökonomie oder der historischen Forschung vorkommen, reicht allerdings auch die Klasse der ARMA-Prozesse nicht aus. Dies deshalb, weil solche Reihen in der Regel nicht mit dem Stationaritätspostulat vereinbar sind, das auch den ARMA-Prozessen zugrundeliegt. Der Hauptgrund dafür liegt darin, daß sie

meistens einen ausgeprägten Trend aufweisen und deshalb nicht um ein konstantes Niveau schwanken, wie das für stationäre Prozesse charakteristisch ist. Um dennoch mit ARMA-Prozessen arbeiten zu können, wurde von Box und Jenkins eine einfache Prozedur vorgeschlagen, die dazu dienen soll, die auf einen Trend zurückzuführende Nicht-Stationarität zu beseitigen. Diese besteht in einer Differenzenbildung:

$$Y_t := (1 - B)^d X_t, \quad d = 1, 2, \dots$$

d.h. die Reihe X_t soll durch Differenzenbildung in eine (stationäre) Reihe Y_t transformiert werden, die dann für die Modellbildung benutzt wird. Dabei gibt der Exponent "d" den "Grad" der Differenzenbildung an. Bei $d = 1$ ist $(1 - B)X_t = X_t - X_{t-1}$. Die hier gewählte vorsichtige Formulierung, daß Stationarität erreicht werden "soll", will darauf hinweisen, daß eine Differenzenbildung natürlich noch keine Garantie für ein stationäres Y_t ist. Allerdings existieren auch Tests zur Überprüfung dieser Hypothese, z.B. der Test von Dickey-Fuller, die aber praktisch eher selten verwendet werden. Hier verläßt man sich mehr auf die Erfahrung, die davon ausgeht, daß sich bei den meisten praktischen Reihen in der Regel durch höchstens zweimalige Differenzierung eine Reihe erzeugen läßt, die "hinreichend" stationär ist. Es sei hier noch kritisch angemerkt, daß die Technik der Differenzenbildung zur Trendelimination in anderen Zusammenhängen bedenklich sein kann, z.B. in der "Langen Wellen"-Forschung. Man kann nämlich zeigen, daß Differenzen eine ganz bestimmte Art von "Hochpaßfilter" sind mit einer sehr unscharfen "Amplitudenfunktion", mit der Konsequenz, daß gerade die interessierenden Bewegungskomponenten ausgefiltert werden. Näheres dazu siehe unten.

Prozesse, die sich nach d-maliger Differenzenbildung als ARMA-Prozesse darstellen lassen, werden als ARIMA(p,d,q)-Prozesse bezeichnet, die sich kurz so schreiben lassen:

$$\Phi(B) (1 - B)^d X_t = \Theta(B) u_t$$

p, d, und q bestimmen die Prozeßgleichung vollständig. Diese lautet z.B. für den ARIMA (1,1,2)-Prozeß:

$$(1 - \Phi_1 B) (1 - B) X_t = u_t - \Theta_1 u_{t-1} - \Theta_2 u_{t-2}$$

oder ausführlich:

$$X_t - (1 + \Phi_1)X_{t-1} + \Phi_1 X_{t-2} = u_t - \Theta_1 u_{t-1} - \Theta_2 u_{t-2}$$

Schließlich sei noch erwähnt, daß für Reihen, die eine saisonale Komponente enthalten, eine Erweiterung notwendig ist. Diese beruht im wesentlichen darauf, daß man sich eine Reihe in einen saisonalen Teil und in einen nicht-saisonalen Teil zerlegt denkt. Für beide Teile wird jeweils ein ARIMA-Modell entwickelt. Beide Modelle werden multiplikativ miteinander verknüpft zu einem Modell für die gesamte Reihe. Für den saisonalen Modellteil ist die Symbolik (P,D,Q) gebräuchlich. D bezeichnet dabei den Grad der saisonalen Differenzenbildung. Bei Monatsreihen z.B. werden für $D=1$ die Differenzen $X_t - X_{t-12}$ gebildet. P bzw. Q bezeichnen den Grad des saisonalen AR- bzw. MA-Polynoms. Das ganze Modell wird bezeichnet mit $ARIMA(p,d,q)*(P,D,Q)_{12}$ für Monatsreihen, mit $ARIMA(p,d,q)*(P,D,Q)_4$ für Quartalsreihen usw. Auf weitere Einzelheiten kann hier nicht näher eingegangen werden.

3. Identifikation, Schätzung und Diagnose von ARIMA-Modellen

Unter der Identifikation eines ARIMA-Modelles versteht man die Festlegung der Ordnungen p , q und d . Einzige Informationsquelle dafür ist die gegebene Zeitreihe, genauer: Kenngrößen, die auf der Basis dieser konkreten Reihe berechnet werden können. Hauptsächlich kommen dafür die Autokorrelationsfunktion und die sog. "partielle" Autokorrelationsfunktion in Frage.

Die Autokorrelationsfunktion wurde oben schon eingeführt. Es kann gezeigt werden, daß diese für einen reinen $MA(q)$ -Prozeß ab dem lag $q+1$ identisch gleich Null ist. Für einen reinen $AR(p)$ -Prozeß verläuft sie jedoch asymptotisch (monoton oder oszillierend) gegen Null. Somit ist diese Funktion prinzipiell geeignet, zwischen diesen beiden Prozeßtypen zu diskriminieren. In der Praxis muß sie allerdings erst geschätzt werden. Dafür existieren geeignete Schätzfunktionen, auf deren Wiedergabe hier verzichtet werden soll. Liegt nun z.B. ein $MA(3)$ -Prozeß vor, dann müßten die geschätzten Autokorrelationen theoretisch ab dem lag 4 identisch gleich Null sein. Praktisch wird das jedoch nicht der Fall sein. Vielmehr werden die geschätzten Autokorrelationen auch ab dem lag 4 noch von Null verschieden sein. Allerdings ist zu erwarten, daß sie ab diesem lag nicht mehr "signifikant" von Null verschieden sind, d.h. für einen $MA(3)$ -Prozeß ist genau mit 3 signifikant von Null abweichenden Autokorrelationen zu rechnen. Die Entscheidung über den vorliegenden Prozeßtyp wird deshalb mit Hilfe von speziellen Signifikanztests vollzogen. Als Prozeßordnung q wird generell diejenige lag-Zahl

identifiziert, deren zugehöriger Korrelationswert sich noch als signifikant verschieden von Null erweist.

Zur Identifikation eines reinen AR-Prozesses verwendet man die partielle Autokorrelationsfunktion. Diese läßt sich (vom Schätzaspekt her) am einfachsten durch folgende Überlegung verständlich machen. Angenommen, es liege ein reiner AR-Prozeß unbekannter Ordnung p vor. Eine naheliegende Strategie zur Bestimmung von p besteht darin, daß man eine Folge von Regressionen berechnet: zuerst X_t auf X_{t-1} , dann X_t auf X_{t-1} und X_{t-2} usw. Ist z.B. $p=4$, dann ist zu erwarten, daß der 5. und alle weiteren geschätzten Regressionskoeffizienten nicht mehr signifikant von Null verschieden sind. Die Entscheidung über Signifikanz bzw. Nichtsignifikanz dieser Regressionskoeffizienten kann mit dem in der Regressionstheorie üblichen t-Test getroffen werden. Die Folge dieser geschätzten Regressionskoeffizienten ist die geschätzte partielle Autokorrelationsfunktion. Als Ordnung p wird daher die Anzahl der von Null signifikanten Koeffizienten gewählt. Im Gegensatz zu AR-Prozessen bricht die partielle Autokorrelation für MA-Prozesse nicht abrupt ab, sondern geht asymptotisch gegen Null. Auto- und partielle Autokorrelation verhalten sich bezüglich der beiden reinen Prozeßtypen AR und MA sozusagen zueinander dual.

Wesentlich schwieriger gestaltet sich die Identifikation von ARMA-Prozessen, da sich für gemischte Prozesse weder für die Autokorrelation noch für die partielle Autokorrelation ein charakteristisches Verlaufsmuster angeben läßt. Beide Korrelationen gehen asymptotisch gegen Null. Deswegen wurden für diesen Fall zusätzliche Identifikationstechniken- bzw. Kriterien entwickelt wie z.B. die Vektorkorrelation, das AIC (Akaike)- und das BIC (Bayes)-Kriterium usw. Darauf kann hier nicht eingegangen werden.

Die Festlegung von d macht in der Regel weniger Schwierigkeiten, meistens genügt $d=1$ oder $d=2$. Generell sollte d so gewählt werden, daß die geschätzte Autokorrelationsfunktion "rasch" gegen Null abfällt (entweder monoton oder oszillierend).

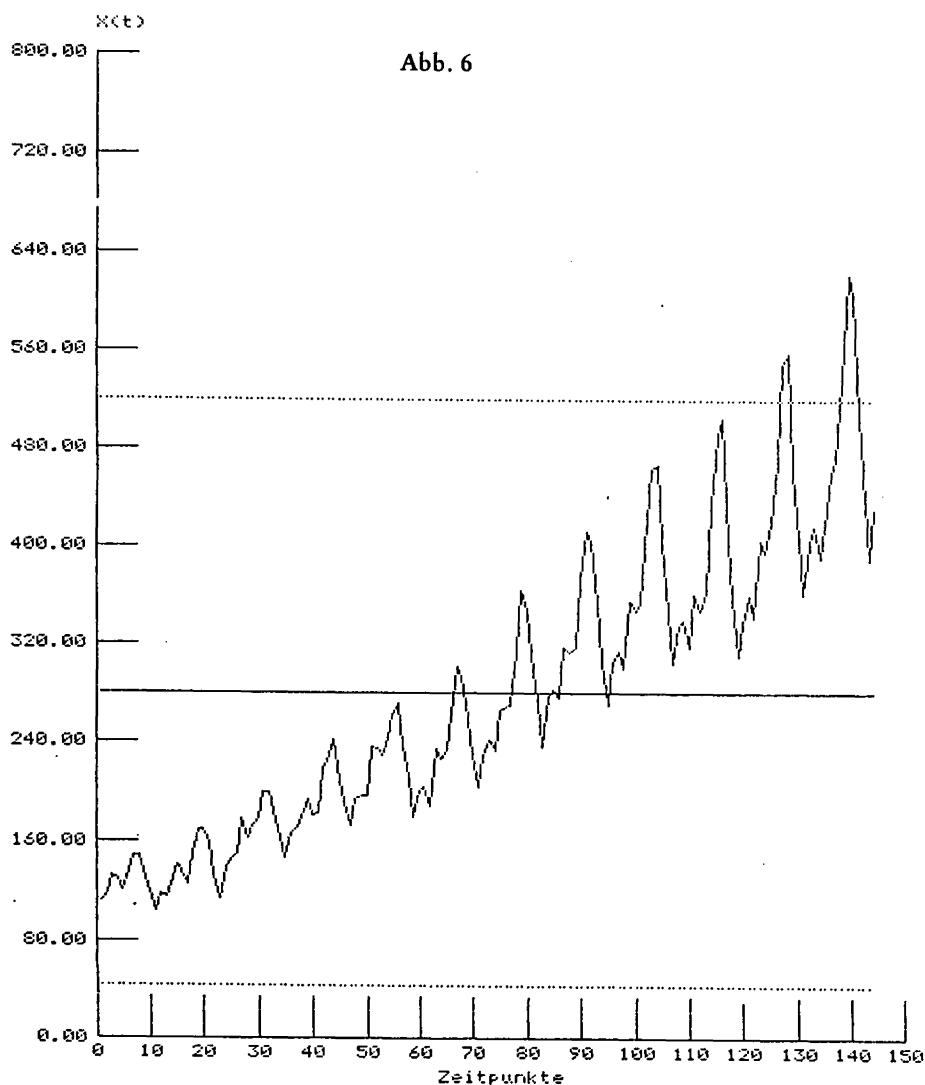
Die Identifikation eines ARIMA-Prozesses kann als ein Mustererkennungsproblem angesehen werden, das eigentlich viel Erfahrung und Finger-spitzengefühl voraussetzt. Allerdings existieren heute auch Software-Pakete (auch für PC) mit automatischen Identifikationsprozeduren wie z.B. AUTOBOX PLUS, so daß auch ein Nicht-Experte mühelos ARIMA-Modelle identifizieren kann.

Nach der Identifikation sind die unbekannten Parameter des ARMA-Prozesses zu schätzen. Dabei geht es im Prinzip darum, diese so zu bestimmen,

daß sich ein möglichst guter "fit" von Modell (genauer: den vom Modell "vorhergesagten" Reihenwerten) und den tatsächlichen Reihenwerten ergibt. Dazu wird entweder die Methode der kleinsten Quadrate verwendet oder die sog. "Maximum-Likelihood"-Methode. Die Likelihood-Funktion kann als Plausibilitätsmaß dafür interpretiert werden, daß der zugrundeliegende Prozeß gerade die vorliegende Zeitreihe erzeugt hat. Es liegt deshalb nahe, die unbekannten Prozeßparameter so zu wählen, daß diese Plausibilität maximal wird. Auf Einzelheiten muß wieder verzichtet werden. Es sei nur noch erwähnt, daß sich ein Anwender um diese ziemlich komplizierten Schätzprozeduren nicht zu kümmern braucht, da in allen einschlägigen Software-Paketen der Schätzvorgang vollständig automatisiert ist.

Als letzter Schritt der Modellentwicklung ist schließlich noch die Modell-diagnose zu nennen. Dabei geht es im wesentlichen um Residuenanalyse. Nach Modellkonzeption sollten die Residuen (die als Differenzen zwischen Modell-Reihenwerten und den tatsächlichen Reihenwerten definiert sind) als Realisation eines white-noise-Prozesses interpretierbar sein, was gleichbedeutend damit ist, daß ihre Autokorrelationen nicht signifikant von Null verschieden sein dürfen. Dies kann mit Hilfe spezieller Tests (z.B. dem Ljung-Box-Test) überprüft werden. Können die Residuen nicht in dieser Weise interpretiert werden, so ist das ein Indiz für eine fehlerhafte Modellspezifikation, was Anlaß für eine erneute Identifikation geben kann. Bei der automatischen Identifikation werden die einzelnen Schritte: Identifikation-Schätzung-Diagnose iterativ so oft durchlaufen, bis sich eine "befriedigende" Diagnose ergibt.

Abschließend sei hier noch eine grundsätzliche Bemerkung angefügt. Der oben skizzierte Ansatz unterstellt, daß sich eine gegebene Zeitreihe als eine Realisation eines ARIMA-Prozesses darstellen läßt. Praktische Reihen sind jedoch häufig "verzerrt" durch "Ausreißer" und/oder "Strukturbrüche" (wie z.B. plötzliche Niveauverschiebungen). Letztere können vorübergehend ("transient") sein oder auch (ab einem bestimmten Zeitpunkt) "permanent". Solche Effekte kann man insbesondere bei historischen Zeitreihen beobachten. Bei der ARIMA-Modellierung ist zu beachten, daß diese Effekte einen starken Einfluß auf die Schätzung der Korrelationsfunktionen haben können, mit der Konsequenz, daß die resultierende Identifikation möglicherweise überwiegend ein Resultat von Reihen-"Irregularitäten" ist. Da diese sich aber nicht unbedingt in der Zukunft wiederholen, kann die Prognosequalität des identifizierten Modelles davon negativ tangiert sein. Allerdings sind diese Effekte einer statistischen Definition und Analyse zugänglich. Mit Hilfe der sog. "Interventionsanalyse" können sie bei der Modellbildung berücksichtigt

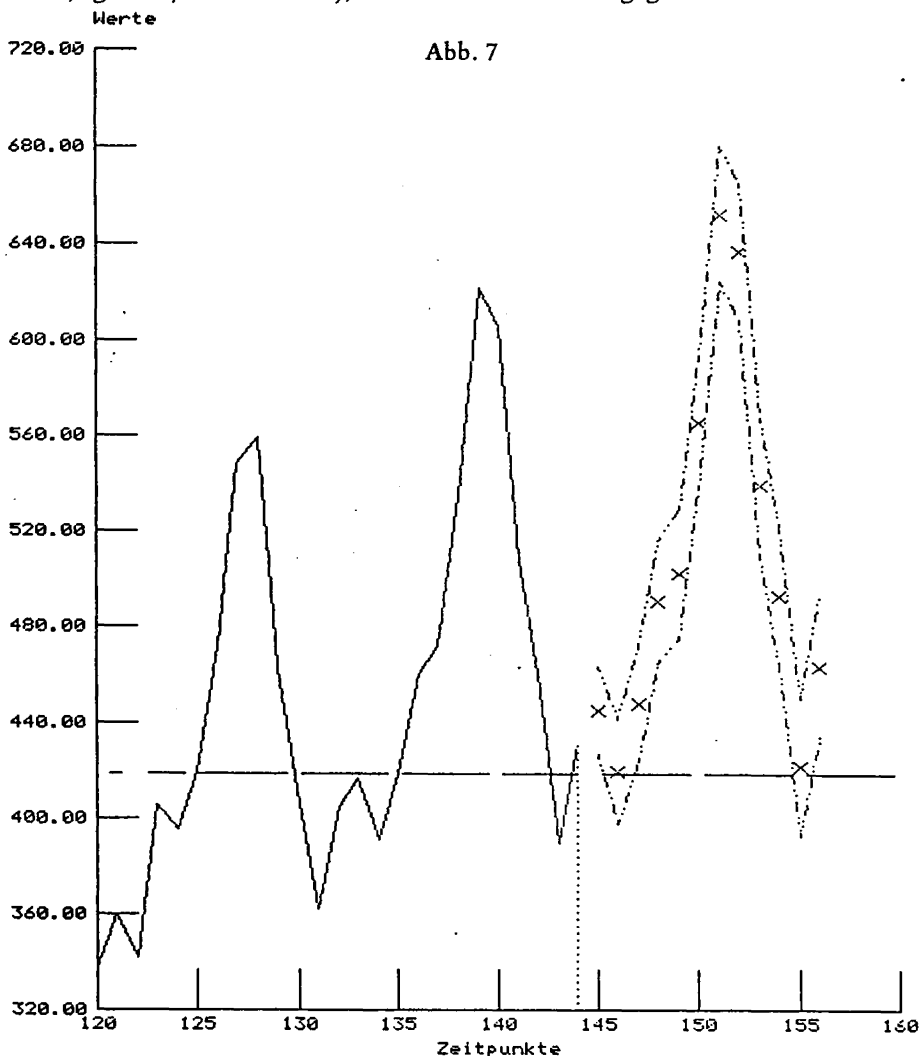


werden. Die angedeuteten Probleme können damit vermieden, zumindest aber in ihren Konsequenzen stark vermindert werden. Im AUTOBOX PLUS z.B. wird eine Interventionsanalyse routinemäßig durchgeführt.

4. Prognose mit Hilfe von ARIMA-Modellen

ARIMA-Zeitreihenmodelle werden vor allem erstellt, um Prognosen durchzuführen. Sie werden jedoch auch bei der Identifikation dynamischer

Beziehungen zwischen zwei oder mehr Zeitreihen benötigt. Aus der identifizierten und geschätzten Prozeßgleichung läßt sich relativ einfach eine Prognosegleichung ableiten. Eine eventuell vorgenommene Differenzbildung muß für die Prognose durch Summenbildung wieder rückgängig gemacht werden. Einzelheiten können wiederum nicht besprochen werden. Stattdessen soll abschließend ein Beispiel für die konkrete Modellierung und Prognose einer Reihe gegeben werden. Verwendet wird dazu die Reihe AIRLINE (Anzahl der Passagiere, monatlich, von Januar 1949 bis Dezember 1960, vgl. Box/Jenkins S.531), die nachstehend wiedergegeben ist:



Die automatische Identifikation mit AUTOBOX PLUS (mit Interventionsanalyse) ergibt nach Parameterschätzung folgendes Modell:

$$(1 - .55268B - .25479B^2)(1 - B^{12})X_t + .074999 = u_t,$$

d.h. einen ARIMA $(2,0,0) \cdot (0,1,0)_{12}$ - Prozeß. Die Prognosewerte für 12 Monate (also bis Dezember 1961) sind in der nächsten Graphik zu finden, wobei die Reihe ab Dezember 1958 wiedergegeben ist:

Die gestrichelten Linien bedeuten die untere bzw. obere 95%- Konfidenzgrenze für die prognostizierten Werte. Diese Intervallprognosen beruhen auf der Annahme, daß der white-noise Prozeß u_t normal verteilt ist.

5. Multivariate ARIMA(Transferfunktionen)-Modelle

Bisher wurden ausschließlich univariate Zeitreihenmodelle betrachtet. Ein nicht unwesentlicher Fortschritt gegenüber dem traditionellen Komponentenmodell besteht darin, daß bei einer Modellierung auf der Grundlage stochastischer Prozesse multivariate Zusammenhänge erfaßt werden können. Eine wichtige Klasse von Modellen stellen die sog. "Transferfunktionenmodelle" dar. Dabei wird eine Reihe als abhängig von mehreren anderen betrachtet. Damit lassen sich z.B. substanzwissenschaftlich formulierte Abhängigkeiten empirisch testen. Existieren solche Abhängigkeiten, dann ist zu erwarten, daß man eine Reihe (nämlich die "abhängige") "besser" prognostizieren kann, wenn man "erklärende" Variablen hinzuzieht, als wenn diese Reihe nur "aus sich selbst" (also univariat) prognostiziert wird. Außerdem ist es möglich, dabei "lead-lag"- Beziehungen aufzuspüren. Beides kann unter bestimmten Bedingungen als Indikator für Kausalbeziehungen interpretiert werden. Diese Möglichkeiten erscheinen auch insbesondere für historische Fragestellungen interessant zu sein. Bisher sind allerdings keine einschlägigen Anwendungen bekannt geworden, was auch damit zusammenhängen mag, daß die Theorie der Transferfunktionenmodelle wesentlich komplizierter ist als diejenige der univariaten ARIMA-Modelle. Aus diesem Grund können hier auch nur einige grundlegende Zusammenhänge dargestellt werden. Dabei soll vom einfachsten Fall mit nur einer erklärenden Variablen ausgegangen werden. Die abhängige Variable sei mit Y_t und die unabhängige mit X_t bezeichnet. Das Transferfunktionsmodell hat dann folgende Gestalt:

$$Y_t = \frac{\alpha(B)}{\delta(B)} X_{t-b} + N_t,$$

wobei der noise-Prozeß nicht unbedingt white-noise, sondern im allgemeinen selbst ein ARIMA(p,d,q)-Prozeß ist:

$$\Phi(B)N_t = \Theta(B)u_t$$

$\omega(B)$ und $\delta(B)$ sind Polynome in B und der Quotient $\omega(B)/\delta(B)$ bezeichnet die "Transferfunktion" zwischen der Input-Variablen X_t und der Output-Variablen Y_t . b ist die sog."reine" Verzögerung zwischen In- und Output. Diese wird nicht vorgegeben, sondern aus den Daten geschätzt.

Das gesamte Modell lautet somit:

$$Y_t = \frac{\omega(B)}{\delta(B)} X_{t-b} + \frac{\theta(B)}{\phi(B)} u_t \quad .$$

Dieses Modell kann erweitert werden durch Einbeziehung weiterer erklärender Variablen. Offensichtlich besteht eine Analogie zwischen Transferfunktions- und Regressionsmodellen. Auch hier verläuft die Modellbildung in einem dreistufigen iterativen Prozeß von Identifikation, Schätzung und Diagnose, wobei die einzelnen Schritte natürlich komplizierter als bei den univariaten Modellen sind. Das sollte aber interessierte Anwender nicht abschrecken, z.B.ist mit AUTOBOX PLUS eine automatische Transferfunktionen-Modellierung möglich.

III. Grundzüge der Filtertheorie

1. Vorbemerkungen

Filtermethoden gehören seit jeher zu den wichtigsten Werkzeugen der Zeitreihenanalyse. Daß sie auch für Fragestellungen in den Geschichtswissenschaften von Bedeutung sein können, kann man am Beispiel der "langen Wellen"-Forschung sehen. Dabei wird in starkem Maß von Filtern Gebrauch gemacht, meistens in Form von gleitenden Durchschnitten. Wie oben ausgeführt wurde, können schon einfache gleitende Durchschnitte als Filter interpretiert werden. Die Begründung dafür ergibt sich aus der Filtertheorie, deren Grundzüge nachstehend skizziert werde. Wiederum wird auf Ableitungen und Details verzichtet.

2. FIR-und IIR-Filter

Unter einem "Filter" kann allgemein ein Algorithmus zur Transformation einer Zeitreihe verstanden werden. Im Rahmen filtertheoretischer Überle-

gungen spricht man statt von Zeitreihen auch häufig von "Signalen". Formal kann der Zusammenhang zwischen der zu filternden Reihe X_t (dem "Filter-Input") und der gefilterten Reihe (dem "Filter-Output") Y_t folgendermaßen dargestellt werden:

$$Y_t = F(X_t).$$

Für theoretische Zwecke ist es oft zweckmäßig, wenn angenommen wird, daß X_t ein Signal unendlicher Länge ist. Die bei der Filterung praktischer Reihen auftretenden Probleme sind allerdings häufig auf deren Endlichkeit zurückzuführen.

Als theoretischer Hintergrund für die hier darzustellenden Filter eignet sich die Theorie der diskreten, linearen, zeitinvarianten Systeme. Dabei wird unter einem "System" jede Vorschrift verstanden, die ein Input-Signal in ein Output-Signal überführt. Linear und zeitinvariant ist ein System, wenn diese Vorschrift durch eine lineare Gleichung mit konstanten Koeffizienten darstellbar ist. So ist z.B. ein einfacher gleitender Durchschnitt mit $2m+1$ Gliedern

$$Y_t = \frac{1}{2m+1} \sum_{t=-m}^m X_t$$

ein zeitinvariantes System. Allgemein kann man für ein lineares System schreiben:

$$Y_t = \sum_{k=-\infty}^{\infty} h_k X_{t-k}, \quad t = \dots, -2, -1, 0, 1, 2, \dots$$

wobei die Folge h_k die sog. "Impuls-Antwortfunktion" des Systems (Filters) ist. Diese ist der Output eines Filters, wenn als Input die sog. "diskrete Impulsfunktion" verwendet wird. Darunter ist ein Signal zu verstehen, das für $t=0$ gleich Eins und sonst gleich Null ist.

Mit Hilfe der Impuls-Antwortfunktion (IAF) können Filter grundsätzlich in zwei Klassen eingeteilt werden: solche mit endlicher IAF und solche mit unendlicher IAF. Die ersteren werden auch als FIR-Filter (Finite-Impulse-Response-Filter) und die letzteren als IIR-Filter (Infinite-Impulse-Response-Filter) bezeichnet. Offensichtlich sind gleitende Durchschnitte spezielle FIR-Filter.

Die Hauptaufgabe der Filtertheorie besteht darin, Methoden zur Konstruktion von Filtern zu entwickeln, die gewissen a priori spezifizierten Kriterien entsprechen. Dabei spielt die sog. "Transferfunktion" eines Filters die

entscheidende Rolle. Sie gibt, grob gesagt an, wie der Filter unterschiedliche Schwingungskomponenten einer Zeitreihe jeweils gewichtet (dazu unten mehr). Jede Zeitreihe läßt sich nämlich, auch wenn sie nicht "zyklisch" aussieht, als Kompositum einer Vielzahl von Schwingungen ("harmonische" Wellen) unterschiedlicher "Frequenzen" auffassen. Dabei ist die Frequenz der Reziprokwert der Schwingungsdauer ("Periode"), die wiederum die Anzahl der Zeiteinheiten darstellt, die eine komplette Schwingung benötigt. Die kleinste beobachtbare Schwingungsdauer beträgt (bei diskreten Reihen) zwei Zeiteinheiten. Die höchste Frequenz ist also $f = 1/2$. Ein linearer Trend läßt sich als eine Welle mit "unendlich langer" Periodendauer konzipieren, die entsprechende Frequenz ist also Null.

Für IIR-Filter läßt sich die Beziehung zwischen Filter-Input und Filter-Output in Form einer linearen Differenzengleichung

$$y_t + a_1 y_{t-1} + \dots + a_M y_{t-M} = b_0 x_t + b_1 x_{t-1} + \dots + b_N x_{t-N}$$

darstellen. Der Filter ist vollständig bestimmt, wenn die Koeffizienten a_1, \dots, a_M und b_1, \dots, b_M bestimmt sind (für FIR-Filter gilt $a_1 = \dots = a_M = 0$). M und N bezeichnen die Ordnung eines Filters. Seine Transferfunktion ist gegeben durch:

$$T(\lambda) = \frac{\sum_{r=0}^N b_r e^{-i\lambda r}}{\sum_{s=0}^M a_s e^{-i\lambda s}}, \quad |\lambda| \leq \pi$$

Offensichtlich hat diese Funktion für FIR-Filter nur Nullstellen, während sie für IIR-Filter sowohl Nullstellen als auch Pole besitzt. (Unter den Polen einer rationalen Funktion sind die Nullstellen des Nenners zu verstehen. An diesen Stellen nimmt die Funktion einen unendlich großen Wert an. Zum Beispiel hat die Funktion $f(z) = (z + 2)/(z-1)$ eine Nullstelle bei $z = -2$ und einen Pol bei $z = 1$). Die Transferfunktion ist im allgemeinen komplexwertig und kann deshalb geschrieben werden als:

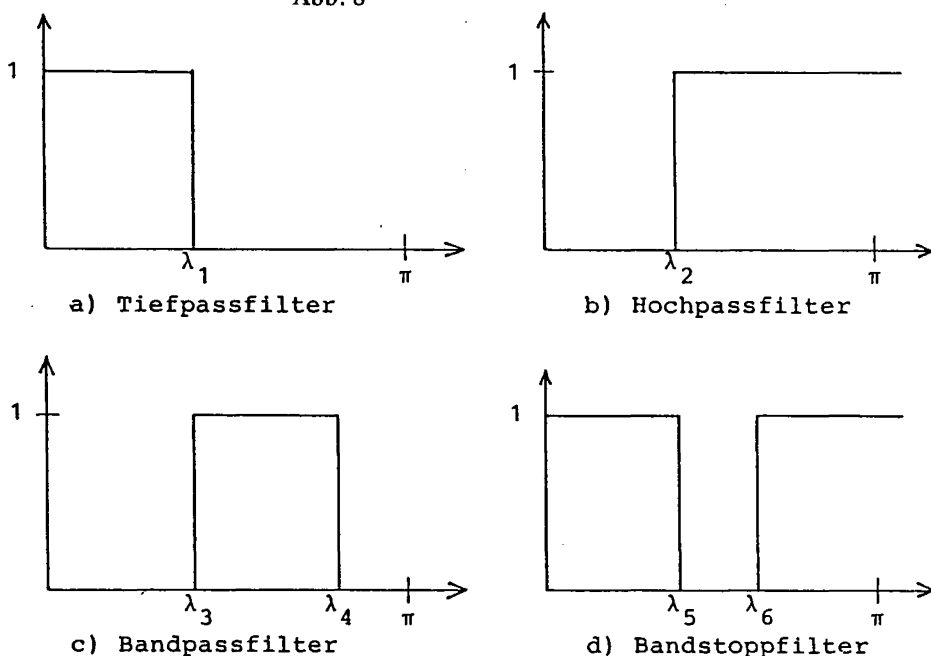
$$T(\lambda) = |T(\lambda)| e^{i\Phi(\lambda)}.$$

Dabei ist $|T(\lambda)| = A(\lambda)$ die sog. "Amplitudenfunktion" und $\Phi(\lambda)$ die sog. "Phasenfunktion" des Filters. Aus der Amplitudenfunktion lassen sich die "Ausblendeigenschaften" eines Filters ermitteln, d.h. die Frage beantwor-

ten, wie die "power" eines Signals, die einzelnen Frequenzen (bzw. Frequenzintervallen) zuzuordnen ist, vom Filter verändert (d.h. abgeschwächt oder eliminiert oder verstärkt) wird. Ein wesentlicher Aspekt beim Filter-Design besteht gerade darin, Filter zu entwerfen, die eine vorgegebene Amplitudenfunktion "realisieren", die nach substanzwissenschaftlichen Gesichtspunkten spezifiziert wurde. Dabei wird in der Regel das Ziel verfolgt, Informationen, die in einem Signal enthalten sind, zu extrahieren, so daß sie quasi "isoliert" auswertbar sind (vgl. dazu das oben skizzierte Problem der Extraktion von "langen Wellen"), oder sie auch zu eliminieren.

Grundlegend und praktisch am wichtigsten sind die vier Filtertypen: Tiefpaß-, Hochpaß-, Bandpaß- und Bandstoppfilter. Deren (ideale) Amplitudenfunktionen folgen nachstehend (dabei wird nur die positive Frequenzachse wiedergegeben, da sie symmetrisch zur Ordinate sind):

Abb. 8



Mit einem Hochpaß z.B. läßt sich eine Trendelimination durchführen, d.h. es werden alle Schwingungskomponenten mit einer Frequenz $\lambda < \lambda_2$ aus der Reihe eliminiert. Dabei ist die Frequenz λ_2 so festzulegen, daß sie mit der jeweiligen substanzwissenschaftlich fundierten Definition von "Trend" korrespondiert. Das gleiche gilt für die Frequenz λ_1 beim Tiefpaß, der einen Trend

extrahiert oder für die Frequenzen λ_3 und λ_4 beim Bandpaß, die z.B. eine "lange Welle" definieren, während die Frequenzen λ_5 und λ_6 beim Bandstopp eine zu eliminierende Komponente definieren. Man sieht, daß hier die Möglichkeit besteht, die einzelnen Filter so zu entwerfen, daß sie einzelne Komponenten einer Reihe "problembezogen" erfassen. Eine derartige Möglichkeit besteht im Rahmen des traditionellen Komponentenmodelles und mit den Werkzeugen der traditionellen Zeitreihenanalyse nicht (vgl. dazu die Ausführungen unter 1.3).

Da obige Amplitudenverläufe als "ideal" bezeichnet wurden, steht zu vermuten, daß reale Filter die dargestellten "exakten" Amplituden nicht unbedingt aufweisen. Unter "exakt" sei hier verstanden, daß die Amplituden in den Paß- bzw. Stoppbändern genau den Wert Eins bzw. Null annehmen. Eine Zielsetzung bei der Filterkonstruktion besteht deshalb in einer möglichst guten Approximation der idealen Amplitudenfunktionen.

Neben der Amplitudenfunktion ist die Phasenfunktion eine wichtige Kenngröße eines Filters. Sie liefert Informationen darüber, mit welcher Verzögerung der Filter-Output dem Filter-Input folgt. Es kann gezeigt werden, daß FIR-Filter eine lineare, (und damit als Spezialfall möglicherweise auch eine Nullphase), IIR-Filter jedoch eine nicht-lineare Phase besitzen.

Es existiert eine Vielfalt von Filter-Designmethoden sowohl für FIR-als auch für IIR-Filter, die hier jedoch nicht einmal ansatzweise besprochen werden können. Der interessierte Leser sei dazu auf Stier (1978, 1990) verwiesen. Vergleichend sei hier lediglich auf einige wichtige charakteristische Unterschiede bzw. Besonderheiten hingewiesen:

a) Mit IIR-Filter lassen sich schon bei niedriger Filterordnung sehr "scharfe" (praktisch "exakte") Amplitudenfunktionen realisieren. Dies ist mit FIR-Filter nicht möglich. Scharfe Amplituden erfordern bei diesem Filtertyp sehr hohe Ordnungen, zu deren Realisierung meistens mehrere hundert Datenpunkte erforderlich sind.

b) Exakte Nullphasen sind nur mit FIR-Filter möglich. Dabei ist aber eine "symmetrische" Implementierung Voraussetzung. "Asymmetrische" Implementierungen erzeugen eine lineare Phase (d.h. jeder Output-Wert ist um den gleichen Betrag gegenüber seinem Inputwert verschoben). Allerdings kann mit IIR-Filter manchmal eine "quasi"- Nullphase erzielt werden.

c) Zwischen der Schärfe der Amplitudenfunktion und der Größe der Phasenverschiebung besteht grundsätzlich eine Antinomie: je schärfer die Amplitude, desto größer ist die Phase.

d) IIR-Filter erzeugen absolut randstabile Filter-Outputs.

3. Filtern im Frequenzbereich

Beim Entwurf von FIR- und IIR-Filter wird sowohl im Frequenz- als auch im Zeitbereich operiert. In der Regel wird die Transferfunktion im Frequenzbereich konstruiert. Daraus werden die Filterkoeffizienten bestimmt und die Filterung im Zeitbereich durchgeführt (via Differenzengleichung). Idealerweise sollte ein Filter die folgenden drei Eigenschaften aufweisen:

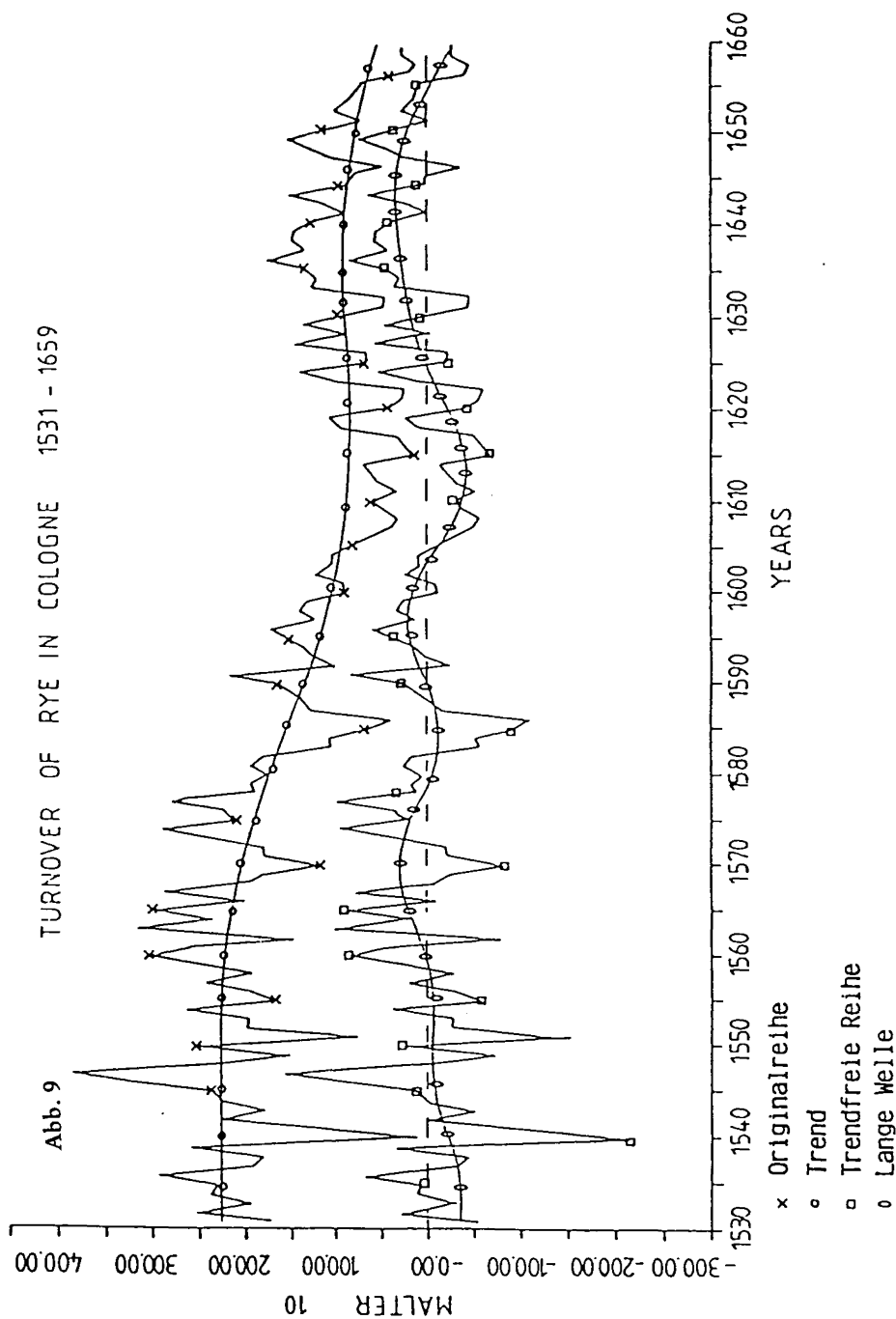
- a) exakte Amplitudenfunktion
- b) exakte Nullphase
- c) absolute Randstabilität.

Man kann sich leicht überlegen, daß weder mit FIR- noch mit IIR-Filter alle drei Desiderata simultan erfüllbar sind. Für manche Analysen erweisen sich aber die Eigenschaften a) und b) als fast unverzichtbar. Dies trifft z.B. auf Probleme zu, wie sie bei der "langen Wellen"-Forschung auftreten. Das "Aufspüren" und gegebenenfalls Extrahieren solcher "Wellen" erfordert Filter, die Postulat a) erfüllen. Da sie als historische Abläufe interpretierbar sein müssen, kommt "Datierungen" ein besonderes Gewicht zu. Deshalb ist auch b) in diesem Kontext unverzichtbar.

Neben den oben erwähnten Filter-Design-Möglichkeiten gibt es jedoch eine dritte, bei der sowohl a) als auch b) erfüllbar sind. Diese ist dadurch charakterisiert, daß auch der Filtervorgang selber im Frequenzbereich stattfindet. Dazu ist es notwendig, die Daten in diesen Bereich zu transformieren, was mit der "diskreten Fourier-Transformation" möglich ist. Die Amplitudenfunktion läßt sich direkt im Frequenzbereich exakt vorgeben. Allerdings ist mit dieser Filtermethode kein absolut randstabiler Output zu erzielen. Auch gibt es dabei ein "Randproblem", das hier jedoch nicht besprochen werden kann. Stattdessen sei abschließend ein Beispiel betrachtet.

Dazu wird die Reihe "Umsatz von Roggen in Köln" (1531-1659) verwendet. Als "Trend" werden alle Bewegungen betrachtet, deren Periode länger als 60 Jahre ist. Zu seiner Extraktion wurde deshalb ein Tiefpaß mit $\lambda_1 = 2\pi/60$ benutzt. Zur Trendelimination diente ein dazu komplementärer Hochpaß. Zur Extraktion evtl. vorhandener "langer Wellen" wurde ein Bandpaß mit $\lambda_3 = 2\pi/60$ und $\lambda_4 = 2\pi/20$ verwendet, da alle Oszillationen mit einer Periodenlänge zwischen 20 und 60 Jahren als "lange Wellen" definiert wurden. Die erzielten Resultate sind in den folgenden Abbildungen wiedergegeben.

Es ist zu betonen, daß beim Filtern keinerlei a priori-Verlaufsformen spezifiziert werden müssen wie beim traditionellen Komponentenmodell. Es ist lediglich erforderlich, daß die interessierenden Komponenten im Frequenz-



bereich nach substanzwissenschaftlichen Gesichtspunkten definiert werden. Wie sie im Zeitbereich verlaufen ist dabei unerheblich. Dieser Verlauf ist sozusagen erst a posteriori, d.h. als Resultat der Filterung, feststellbar.

4. Filter-Design und Spektralanalyse.

Immer wieder werden in Diskussionen Filter-Design und Spektralanalyse miteinander in Verbindung gebracht. Nicht selten wird beides - fälschlicherweise - gleichgesetzt. Da diese Vorstellung anscheinend nur schwer auszumerzen ist, sei hier speziell auf einen Vergleich dieser beiden Gebiete eingegangen.

Gemeinsam ist beiden, daß sie es mit dem Frequenzbereich zu tun haben. Das ist aber auch schon alles. Die Fragestellung der Spektralanalyse läßt sich kurz so charakterisieren:

Ausgangspunkt ist die Theorie der schwach stationären Prozesse. Im Rahmen dieser Theorie wird gezeigt, daß für jeden schwach stationären Prozeß unter bestimmten Bedingungen ein sog. "Spektrum" existiert. Falls es "normiert" wird, ist auch die Bezeichnung "Spektraldichte" gebräuchlich. Die Normierung hat zur Folge, daß die Fläche, welche diese (positive) Funktion mit der Frequenzachse einschließt, gleich Eins ist. Die Spektraldichte hat formal die gleichen Eigenschaften wie die Dichtefunktion einer Zufallsvariablen. Inhaltlich kann sie so interpretiert werden, daß sie den Frequenzgehalt eines Prozesses repräsentiert, genauer: seine Verteilung auf einzelne Frequenzbänder. Da das Integral über das Spektrum gleich der Varianz des stationären Prozesses ist, kann auch davon gesprochen werden, daß das Spektrum (bzw. die Spektraldichte) zeigt, welche Frequenzbänder den zeitlichen Verlauf eines stationären Prozesses hauptsächlich bestimmen. Dominieren z.B. die hohen Frequenzen, dann ist dieser Verlauf ziemlich unruhig, während bei Dominanz von niedrigen Frequenzen ein eher "glatter" Verlauf zu erwarten ist. Daneben kann es bei einzelnen Frequenzen "peaks" in einem Spektrum geben, die auf die Existenz einer zyklischen Komponente dieser Frequenz im Prozeß hinweist.

Im allgemeinen ist das Spektrum bzw. die Spektraldichte unbekannt. Liegt jedoch eine Realisation eines stationären Prozesses vor, d.h. eine (endliche) Zeitreihe, dann kann es bzw. sie geschätzt werden. Dazu gibt es mehrere Möglichkeiten, auf die hier nicht eingegangen werden kann.

Damit läßt sich der Unterschied zwischen Filter-Design und Spektralanalyse wie folgt zusammenfassen:

Bei der Spektralanalyse wird versucht, eine unbekannte Funktion auf der

Grundlage einer Stichprobe zu schätzen. Es liegt also ein Inferenzproblem vor, allerdings kein in der Statistik übliches, da kein(e) Parameter geschätzt wird (werden), sondern eine Funktion. Beim Filter-Design liegt dagegen ein Konstruktionsproblem vor. Werden z.B. IIR-Filter konstruiert, dann ergeben sich die Filter-Koeffizienten nicht via Schätzung aus einer vorliegenden Zeitreihe. Sie werden vielmehr ganz unabhängig von einer konkreten Reihe z.B. nach gewissen Optimalitätsprinzipien bestimmt. Dies hat mit statistischer Inferenz überhaupt nichts zu tun. Somit behandeln beide Gebiete nicht nur völlig verschiedene Fragestellungen, sie sind auch methodisch nicht vergleichbar.

5. Filter-und ARIMA-Modelle

Ein wesentlicher Schritt bei der ARIMA-Modellierung ist die Differenzbildung. Wie man leicht sieht, ist eine Bildung von Differenzen gleichbedeutend mit einer Filterung. Man kann zeigen, daß Differenzen nichts anderes sind als Hochpaßfilter. Leider handelt es sich dabei um ausgesprochen schlechte Hochpaßfilter. Verglichen mit der oben dargestellten Amplituden-Idealform weisen sie vor allem zwei gravierende Nachteile auf:

a) Im Niederfrequenzbereich zeigt die Amplitudenfunktion eines Differenzenfilters einen nur langsam ansteigenden Verlauf, d.h. sie erreicht den Wert Eins erst bei relativ hohen Frequenzen.

b) Im Hochfrequenzbereich nimmt sie Werte an, die wesentlich größer als Eins sind (bei der höchsten Frequenz $\lambda = \pi$ ist $A(\pi) = 2$).

Hinzukommt, daß Differenzenfilter eine Phasenverschiebung zwischen der Originalreihe und der gefilterten Reihe bewirken.

Die Eigenschaft b) führt dazu, daß so gefilterte Reihen "aufgeraut" werden, also einen wesentlich unruhigeren Verlauf zeigen als die Ausgangsreihe. Mit steigender Differenzenordnung wird dieser Effekt noch verstärkt. Eigenschaft a) kann dazu führen, daß gerade die Informationen, für die man sich interessiert, verlorengehen. Diese Gefahr ist insbesondere bei der Suche nach "langen Wellen" gegeben, da solche, falls sie in einer Reihe vorhanden sind, bei sehr niedrigen Frequenzen zu suchen sind. Anders ausgedrückt: die "power", die solchen Wellen entspricht, residiert im Niederfrequenzbereich. Aber gerade dort weisen Differenzenfilter sehr kleine Werte auf, d.h. diese "power" wird durch die Filterung praktisch eliminiert. Ein Hochpaß, der für die vorliegende Fragestellung geeignet ist, müßte in diesem Bereich den Wert Eins annehmen. Nebenbei bemerkt, sind die gleichen Bedenken auch für eine

Suche nach "langen Wellen" auf der Basis von Wachstumsraten anzumelden, da diese approximativ als Differenzenfilter angesehen werden können.

Diese Defekte der Differenzenbildung schließen jedoch eine ARIMA-Modellierung von "langen Wellen" - Phänomenen keineswegs aus. Dazu ist es allerdings erforderlich, die Differenzenbildung durch eine der substanzwissenschaftlichen Fragestellung angemessene Hochpaß- bzw. Bandpaßfilterung (wenn man nur an der "Welle" interessiert ist) zu ersetzen. Deswegen sollte dann auch zutreffender nur von ARMA-Modellierung gesprochen werden. Ein Beispiel soll die Verknüpfung von Filtern und Modellierung veranschaulichen. Dazu wird die Jahresreihe "Gross Fixed Capital Formation in the United Kingdom 1830-1979" verwendet, die nachstehend zusammen mit ihrer "langen Welle" wiedergegeben ist. Diese ergibt sich aus der Originalreihe mit Hilfe eines Bandpasses, der alle zyklischen Bewegungen mit einer Periodenlänge von 20 bis 60 Jahren konserviert.

Versucht man die offensichtlich nicht-stationäre Originalreihe in der üblichen Weise zu modellieren, dann findet man als bestes Modell den ARIMA(0,1,4)-Prozeß:

$$X_t - X_{t-1} = u_t + 0.47u_{t-1} + 0.15u_{t-2} + 0.24u_{t-3} + 0.19u_{t-4}$$

Dieser Prozeß läßt keinerlei zyklische Struktur erkennen (dazu müßte ein AR-Teil mindestens der Ordnung zwei vorhanden sein). Die "Welle" läßt sich sehr gut durch den ARMA(2,1)- Prozeß

$$X_t - 1.889X_{t-1} + 0.943X_{t-2} = u_t - 0.45u_{t-1}$$

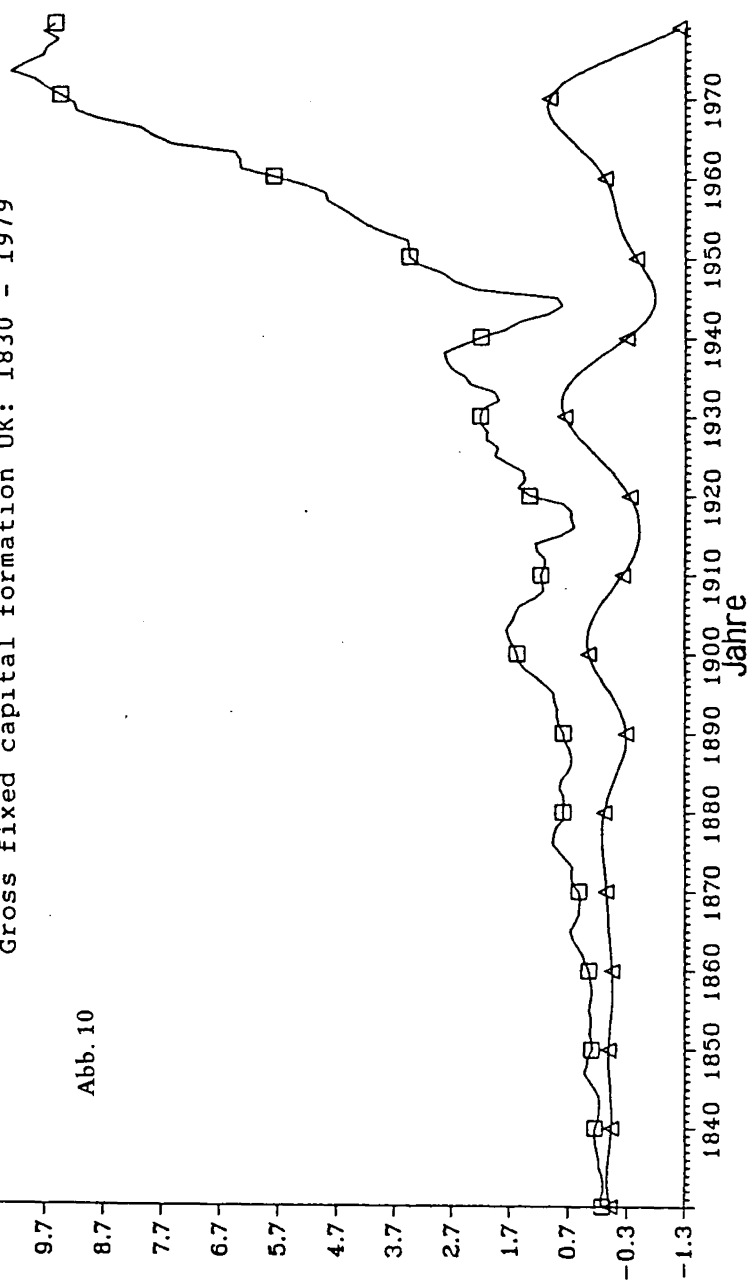
modellieren. Zu dieser Modellierung sei noch eine kurze Bemerkung angefügt: Da die "Welle" einen sehr glatten Verlauf zeigt (alle hochfrequenten Schwingungen wurden ja durch den exakten Bandpaß eliminiert), verbleiben nach der ARMA-Modellierung keine Residuen, die als weißes Rauschen interpretiert werden könnten - entgegen der "Box-Jenkins-Philosophie". Die betragsmäßig sehr kleinen Residuen zeichnen vielmehr den Verlauf der "Welle" nach. Diese Inkonsistenz mit dem Box-Jenkins-Ansatz läßt sich allerdings vermeiden, wenn man die Amplitudenfunktion des Bandpasses im Stoppband des Hochfrequenzbereiches nicht exakt auf Null setzt. Damit bleibt noch etwas "noise" in der "Welle" erhalten, und die erwähnte Inkonsistenz verschwindet.

b₁*

MIO. POUNDS

Gross fixed capital formation UK: 1830 - 1979

Abb. 10



□ = Original
 △ = Lange Welle

Offensichtlich liegt es nahe, den hier vorgeführten Ansatz zu einem multivariaten zu verallgemeinern. Mit Hilfe von Transferfunktionen-Modellen könnten Abhängigkeiten von "langen Wellen" untersucht werden, was für historische Fragestellungen von Interesse sein dürfte. Das scheint aber noch Neuland zu sein. Bisher sind dazu keine Arbeiten bekannt geworden.

Literatur:

- Box, G.E.P., Jenkins, G.M.: Time Series Analysis: Forecast and Control, San Francisco 1976.
- Cryer, J.D.: Time Series Analysis, Duxbury Preß, Boston 1986.
- McCleary R., Hay R.A. jr.: Applied Time Series Analysis for the Social Science, Sage Publ. Beverly Hills and London, 1980.
- Metz, R.: Ansätze, Begriffe und Verfahren der Analyse ökonomischer Zeitreihen, in: Historical Social Research / Historische Sozialforschung, Vol. 13, No. 3 (1980), S.23-103.
- Metz, R.: Zur empirischen Evidenz "langer Wellen", in: Kyklos, Bd. 37, H. 2 (1984), S.266-290.
- Stier, W.: Konstruktion und Einsatz von Digitalfiltern zur Analyse und Prognose ökonomischer Zeitreihen, Westdeutscher Verlag, Opladen 1978.
- Stier, W.: Saisonbereinigungsverfahren - Ein Überblick, in: OR - SPEKTRUM, 1985, 7, S.63-79.
- Stier, W.: Methoden der Zeitreihenanalyse, Springer-Verlag, erscheint 1991.